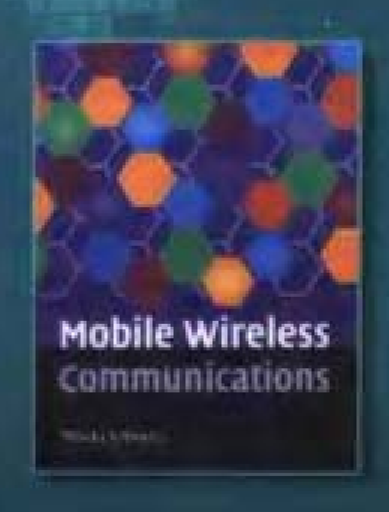
国外电子与通信教材系列

移动无线通信

Mobile Wireless Communications



[美] Mischa Schwartz 著
许希斌 李云洲 译
姚 彦 审校



電子工業出版社
Publishing House of Electronics Industry
http://www.phei.com.cn

移动无线通信

Mobile Wireless Communications

从全球蜂窝电话系统到本地甚至个人网络，无线通信已经成为人们生活中无所不在的一部分。

本书提供了对数字移动无线网络的指导性介绍，书中使用了大量的现实世界的例子来阐明移动无线通信的理论基础。本书从对传播现象的回顾开始，进而探讨了信道分配、调制技术、复接方法以及编码技术等基本内容。书中评述了GSM，IS-136和IS-95系统，并详细讨论了2.5G和3G包交换系统。本书还探讨了性能分析评价和调度技术，并以讲述无线局域网和个人区域网的一章作为结束。本书还提供了许多示例和作业练习，并可为教师免费提供习题解答手册。

这本书是学习无线通信课程的电子工程专业和计算机科学专业学生的理想教材。对于相关行业的工程师而言，同样是一本难得的参考书。

作者简介

Mischa Schwartz 从1974年开始在美国哥伦比亚大学电子工程系任职，现在是Charles Batchelor名誉教授。他编写或合作编写了9本著作，其中包括通信系统和计算机网络方面的畅销书。他现在的研究方向是无线网络。Schwartz教授是美国国家工程学会成员，IEEE会员和前理事，IEEE通信协会前会长，他还是IEEE信息理论分会的前任主席。1983年，他获得了IEEE教育奖。在1984年的IEEE调查中，他是前十位电子工程最佳教育者之一。Schwartz还是美国科学技术研究院的成员，1994年因对计算机通信方面的贡献而获得了纽约市市长奖，2003年因对电信和工程教育的贡献，还获得了日本Okawa奖









责任编辑：谭海平
余 义
责任美编：毛惠庚

ISBN 7-121-02661-9 定价：39.00元

国外电子与通信教材系列

移动无线通信

Mobile Wireless Communications

[美] Mischa Schwartz 著

许希斌 李云洲 译

姚 彦 审校

電子工業出版社

Publishing House of Electronics Industry

北京 · BEIJING

内容简介

这是一本全面介绍移动无线通信基础知识以及最新技术发展状况的教材。

第1章为介绍性内容，讲述移动通信的历史以及后续各章的描述。其后的内容可分为两大部分，第1部分包括第2章到第8章。第2章讲述在无线介质中遇到的传播现象；第3章介绍蜂窝的概念；第4章到第7章讨论功率控制、调制、编码以及接入技术；第8章对第二代系统——GSM，IS-95，IS-136以及D-AMPS进行了详细的讨论。第2部分包括第9章到第12章。第9章探讨了无线系统的性能分析；第10章对第三代系统——W-CDMA、CDMA2000以及GPRS进行了深入的描述；第11章讨论了在蜂窝系统中提出和使用的接入和调度技术。最后全书以对无线局域网（WLAN）和个人区域网（PAN）的讨论作为结束。

本书适合作为大学本科高年级或研究生一年级无线通信的入门课程，也可作为工程师、计算机专家以及其他技术人员的参考材料。

版权贸易合同登记号 图字：01-2005-2686

图书在版编目（CIP）数据

移动无线通信/（美）施瓦茨（Schwartz, M）著．许希斌，李云洲译．－北京：电子工业出版社，2006.9
（国外电子与通信教材系列）
书名原文：Mobile Wireless Communications
ISBN 7-121-02661-9

Ⅰ．移… Ⅱ．①施… ②许… ③李… Ⅲ．移动通信：无线电通信－教材 Ⅳ．TN929.5

中国版本图书馆CIP数据核字（2006）第076727号

责任编辑：谭海平 余 义
印 刷：
装 订：北京牛山世兴印刷厂
出版发行：电子工业出版社
北京市海淀区万寿路173信箱 邮编：100036
经 销：各地新华书店
开 本：787×1092 1/16 印张：23.75 字数：608千字
印 次：2006年9月第1次印刷
定 价：39.00元

凡所购买电子工业出版社的图书有缺损问题，请向购买书店调换；若书店售缺，请与本社发行部联系。联系电话：（010）68279077 邮购电话：（010）88254888

质量投诉请发邮件至zlts@phei.com.cn 服务热线：（010）88258888。

序

2001年7月间，电子工业出版社的领导同志邀请各高校十几位通信领域方面的老师，商量引进国外教材问题。与会同志对出版社提出的计划十分赞同，大家认为，这对我国通信事业、特别是对高等院校通信学科的教学工作会很有好处。

教材建设是高校教学建设的主要内容之一。编写、出版一本好的教材，意味着开设了一门好的课程，甚至可能预示着一个崭新学科的诞生。20世纪40年代MIT林肯实验室出版的一套28本雷达丛书，对近代电子学科、特别是对雷达技术的推动作用，就是一个很好的例子。

我国领导部门对教材建设一直非常重视。20世纪80年代，在原教委教材编审委员会的领导下，汇集了高等院校几百位富有教学经验的专家，编写、出版了一大批教材；很多院校还根据学校的特点和需要，陆续编写了大量的讲义和参考书。这些教材对高校的教学工作发挥了极好的作用。近年来，随着教学改革不断深入和科学技术的飞速进步，有的教材内容已比较陈旧、落后，难以适应教学的要求，特别是在电子学和通信技术发展神速、可以讲是日新月异的今天，如何适应这种情况，更是一个必须认真考虑的问题。解决这个问题，除了依靠高校的老师和专家撰写新的符合要求的教科书外，引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，是会有好处的。

一年多来，电子工业出版社为此做了很多工作。他们成立了一个“国外电子与通信教材系列”项目组，选派了富有经验的业务骨干负责有关工作，收集了230余种通信教材和参考书的详细资料，调来了100余种原版教材样书，依靠由20余位专家组成的出版委员会，从中精选了40多种，内容丰富，覆盖了电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等方面，既可作为通信专业本科生和研究生的教学用书，也可作为有关专业人员的参考材料。此外，这批教材，有的翻译为中文，还有部分教材直接影印出版，以供教师用英语直接授课。希望这些教材的引进和出版对高校通信教学和教材改革能起一定作用。

在这里，我还要感谢参加工作的各位教授、专家、老师与参加翻译、编辑和出版的同志们。各位专家认真负责、严谨细致、不辞辛劳、不怕琐碎和精益求精的态度，充分体现了中国教育工作者和出版工作者的良好美德。

随着我国经济建设的发展和科学技术的不断进步，对高校教学工作会不断提出新的要求和希望。我想，无论如何，要做好引进国外教材的工作，一定要联系我国的实际。教材和学术专著不同，既要注意科学性、学术性，也要重视可读性，要深入浅出，便于读者自学；引进的教材要适应高校教学改革的需要，针对目前一些教材内容较为陈旧的问题，有目的地引进一些先进的和正在发展中的交叉学科的参考书；要与国内出版的教材相配套，安排好出版英文原版教材和翻译教材的比例。我们努力使这套教材能尽量满足上述要求，希望它们能放在学生们的课桌上，发挥一定的作用。

最后，预祝“国外电子与通信教材系列”项目取得成功，为我国电子与通信教学和通信产业的发展培土施肥。也恳切希望读者能对这些书籍的不足之处、特别是翻译中存在的问题，提出意见和建议，以便再版时更正。

吴佑寿

中国工程院院士、清华大学教授

“国外电子与通信教材系列”出版委员会主任

出 版 说 明

进入21世纪以来，我国信息产业在生产和科研方面都大大加快了发展速度，并已成为国民经济发展的支柱产业之一。但是，与世界上其他信息产业发达的国家相比，我国在技术开发、教育培训等方面都还存在着较大的差距。特别是在加入WTO后的今天，我国信息产业面临着国外竞争对手的严峻挑战。

作为我国信息产业的专业科技出版社，我们始终关注着全球电子信息技术的发展方向，始终把引进国外优秀电子与通信信息技术教材和专业书籍放在我们工作的重要位置上。在2000年至2001年间，我社先后从世界著名出版公司引进出版了40余种教材，形成了一套"国外计算机科学教材系列"，在全国高校以及科研部门中受到了欢迎和好评，得到了计算机领域的广大教师与科研工作者的充分肯定。

引进和出版一些国外优秀电子与通信教材，尤其是有选择地引进一批英文原版教材，将有助于我国信息产业培养具有国际竞争能力的技术人才，也将有助于我国国内在电子与通信教学工作中掌握和跟踪国际发展水平。根据国内信息产业的现状、教育部《关于"十五"期间普通高等教育教材建设与改革的意见》的指示精神以及高等院校老师们反映的各种意见，我们决定引进"国外电子与通信教材系列"，并随后开展了大量准备工作。此次引进的国外电子与通信教材均来自国际著名出版商，其中影印教材约占一半。教材内容涉及的学科方向包括电路理论与应用、信号与系统、数字信号处理、微电子、通信系统、电磁场与微波等，其中既有本科专业课程教材，也有研究生课程教材，以适应不同院系、不同专业、不同层次的师生对教材的需求，广大师生可自由选择和自由组合使用。我们还将与国外出版商一起，陆续推出一些教材的教学支持资料，为授课教师提供帮助。

此外，"国外电子与通信教材系列"的引进和出版工作得到了教育部高等教育司的大力支持和帮助，其中的部分引进教材已通过"教育部高等学校电子信息科学与工程类专业教学指导委员会"的审核，并得到教育部高等教育司的批准，纳入了"教育部高等教育司推荐——国外优秀信息科学与技术系列教学用书"。

为做好该系列教材的翻译工作，我们聘请了清华大学、北京大学、北京邮电大学、南京邮电大学、东南大学、西安交通大学、天津大学、西安电子科技大学、电子科技大学、中山大学、哈尔滨工业大学、西南交通大学等著名高校的教授和骨干教师参与教材的翻译和审校工作。许多教授在国内电子与通信专业领域享有较高的声望，具有丰富的教学经验，他们的渊博学识从根本上保证了教材的翻译质量和专业学术方面的严格与准确。我们在此对他们的辛勤工作与贡献表示衷心的感谢。此外，对于编辑的选择，我们达到了专业对口；对于从英文原书中发现的错误，我们通过与作者联络、从网上下载勘误表等方式，逐一进行了修订；同时，我们对审校、排版、印制质量进行了严格把关。

今后，我们将进一步加强同各高校教师的密切关系，努力引进更多的国外优秀教材和教学参考书，为我国电子与通信教材达到世界先进水平而努力。由于我们对国内外电子与通信教育的发展仍存在一些认识上的不足，在选题、翻译、出版等方面的工作中还有许多需要改进的地方，恳请广大师生和读者提出批评及建议。

电子工业出版社

教材出版委员会

译 者 序

自人类社会进入工业文明以来，先后经历了蒸汽机时代、电力时代，到 20 世纪中叶又开始进入了另一个具有革命意义的时代——一个以计算机、通信和信息为支撑的信息网络时代。从 20 世纪 70 年代起，伴随着光纤技术、数字信号处理和超大规模集成电路的突破性发展，有线传输速率已经获得了成百上千倍的增长，到 20 世纪 90 年代末，已经达到几十 T b/s 的速率。与之相对应的无线接入却因为受到频谱资源和移动终端功率的限制无法获得根本性的突破，同时由于无线传输环境的开放性，不可避免地受到随机衰落和干扰的影响，因此无线传输技术成为整个信息传输网络的瓶颈，这也就是所谓的“最后一千米”问题。但是，由于无线移动通信具有无以伦比的便捷性，使得人类对在任何时间、任何地点、与任何人、以任何方式方便快捷地进行信息交流充满了憧憬。在这种需求之下，无线移动通信当之无愧地成为通信领域最大的研究热点之一。在 21 世纪里，无线移动通信无疑必将继续保持迅猛的发展势头，并扩展到社会生活的各个领域，为人们提供更多的便利。

我们所翻译的 Mischa Schwartz 教授的《移动无线通信》一书是 Schwartz 教授在哥伦比亚大学电子工程系任职 30 多年的教学和科研成果的总结，也是一本畅销于全球的、无线移动通信方面的经典著作。本书侧重于阐述无线通信的基本概念、基本技术和无线通信系统基本构成和工作原理。首先，本书详细阐述了无线信号在空中传播的特征，给出了从农村到城郊、城区室外室内不同场景下的信道传播特征，引出了克服信号衰落和信号畸变的分集技术和均衡技术。在链路传输技术方面，本书分别在第 5 章和第 7 章介绍了调制解调技术和信道编译码技术。在系统构成技术方面，本书用了 4 章的篇幅描述了多址技术、蜂窝结构与信道复用技术、接入与调度技术、接纳控制与切换机制，并定量地给出了一些系统性能的分析。而在第 4 章本书进一步介绍了动态信号分配和功率控制两种提升系统性能的方法。就目前运行的无线通信系统而言，以蜂窝为基本架构的蜂窝移动通信系统仍然占据主流，面无线城域网、局域网、个域网也呈现蓬勃发展之势。不管无线通信系统的实际应用如何，为便于阐述无线通信系统的运行原理，本书的重点仍然是介绍数字蜂窝无线通信系统，分别在第 8 章和第 10 章介绍了以电路交换为特征的第二代移动通信系统和引人包交换机制的 2.5/3G 移动通信系统的工作机理。在本书的最后一章，作者介绍了 IEEE 802.11 无线局域网和 IEEE 802.15.1 无线个域网。由于这些系统构成技术与实际系统密切相关，所以本书并没有一章接一章地连续阐述这些技术，而是将这些技术与系统描述、系统性能分析穿插起来介绍，以期读者可以更清晰地理解这些内容。

近年来，无线移动通信技术发展极快，涉及内容也十分广泛，所以很难在一本书中既侧重基本概念介绍，又囊括所有前沿技术和应用。本书的主旨是提供无线移动通信的教学性介绍。它可作为为本科四年级或研究生一年级无线通信课程而设计的入门课本，也可作为工程师、计算机科学家以及其他技术人员的自学教材，并为这一蓬勃发展的领域提供了一个基本的介绍。阅读本书的学生或读者，将能得到无线移动通信方面基础知识的透彻了解，以及对第二代、

第三代蜂窝系统和无线局域网的操作原理的理解。为了加强对引入的各种概念的理解，本书全篇提供了数值和定量的例子。每章的习题都提供了进一步了解该领域知识的途径。同时，书中出现的现有技术文献都有参考资料。作者认为，这些参考资料是讨论的有机部分，当需要的时候能为读者在该领域的技术方面提供深入研究的机会。

本书的前言、第1章、第2章、第5章、第7章由许希斌教授翻译，第3章、第4章、第6章、第8章、第9章、第10章、第11章、第12章由李云洲博士翻译，全书由姚彦教授审校。在本书的翻译过程中电子工业出版社的编辑马岚和余义为本书的出版付出了辛勤的劳动。本实验室的刘杉坚、郑艳丽、辛艳、毛炜、魏铮、苏毅、周婷、马静、张进、许云、张昊、刘涛等同学协助做了大量的工作。借此机会，表示诚挚的谢意。由于水平有限，译文尚有疏漏不当之处，敬请读者不吝指正。

译者

2006年于清华大学

前　　言

即使对于最不经意的观察者，这样的事实也是显而易见的：一场名副其实的电信革命正在发生。无线通信的使用在世界范围内已经得到戏剧性的扩展。蜂窝电话已经随处可见。虽然大多数这样的移动终端仍然主要承载语音业务，然而越来越多的用户已经开始利用它们从事数据和图像业务。无线局域网（例如 Wi-Fi）已经逐渐流行起来，并正在溶入全球化的蜂窝网络。

本书是为大学本科高年级或研究生一年级无线通信课程而设计的入门教材，还可作为工程师、计算机专家以及其他技术人员的自学材料，为这一蓬勃发展的领域提供了一个基本的介绍。本书的读者将能得到无线移动通信方面基础知识的透彻了解，以及对第二代和第三代蜂窝系统和无线局域网的操作原理的理解。为了加强对引入的各种概念的理解，本书全篇都提供了数值和定量的例子。每章的习题都提供了进一步了解该领域知识的途径。同时，书中出现的现有技术文献都有参考资料。作者认为，这些参考资料是讨论的有机部分，当需要的时候能为读者在该领域的技术方面提供深入研究的机会。

第 1 章为介绍性内容，讲述了移动通信的历史以及本书每一章详细的描述。其后的内容分为两大部分，第一部分包括第 2 章到第 8 章，第二部分包括第 9 章到第 12 章。第 2 章到第 7 章介绍了无线移动通信的基础知识，第 8 章对应用这些基本概念的第二代系统——GSM、IS-95、IS-136 以及 D-AMPS 提供了详细的讨论。具体地说，第 2 章处理在无线介质中遇到的传播现象，第 3 章介绍蜂窝的概念。第 4 章到第 7 章讨论了功率控制、调制、编码以及接入技术。这些材料已经在美国哥伦比亚大学一学年的移动通信课程中的第一学期讲授。而第二学期的课程包含本书第二部分的内容以及当前阅读和研究的补充材料。

本书的第二部分提供了更高级的内容。第 9 章深入探讨了无线系统的性能分析，这建立在前面章节简要介绍的一些概念上。第 10 章在第 8 章对第二代系统的讨论基础上对第三代系统——W-CDMA、CDMA2000 以及 GPRS 进行了深入的描述。这一章的重点在采用包交换的数据和多媒体无线通信，还包括了分层结构概念的简要回顾，具体展示了 GPRS 如何与基于 Internet 包交换的网络进行接口。第 11 章讨论了在蜂窝系统中提出和使用的接入和调度技术。本书以对无线局域网（WLAN）和个人区域网（PAN）的讨论作为结束。WLAN 的讨论重点在 Wi-Fi 及其对更高比特率无线局域网的扩展；PAN 的讨论主要集中于蓝牙系统。

因此，本书可以作为无线系统方面的一学期课程或全年的课程。一学期课程中可以像在哥伦比亚大学一样只讲授前 8 章，也可以将第 8 章、第 10 章和第 12 章的材料抽出来作为当前无线系统的例子使用。全年的课程内容可以包括整本书。本书的材料也可以与通信系统课程结合使用，为通信技术在无线领域提供应用。

除了在本科四年级或研究生一年级阶段通常要求具有的基础知识外，本书要求的先验知识只包含基本的概率论和线性代数知识。在通信理论和通信系统方面，本书对读者没有先验知识的要求，例如本书第一部分关于调制和编码的材料是自成一体的。对于在通信系统方面

有先验知识的读者来说,关于调制和编码的章节中的指导性讨论将是一个有用的回顾,是调制和编码理论在无线系统方面的具体应用,从面进一步巩固了读者对该领域知识的理解。

作者希望对在本书写作中共同工作过的很多人的帮助表示感谢。特别地,作者对那些在哥伦比亚大学选修无线课程并使用包含本书材料的原始讲义的学生以及执教该课程的老师表示感谢。作者要特别感谢如下人士的帮助和支持:哥伦比亚大学电子工程系的 Andrew Campbell 教授;曾经在贝尔实验室工作,现在在美国宾西法尼亚州立大学计算机系的 Tom LaPorta 教授以及 IBM 的 Mahmoud Naghshineh 博士。同时,作者也在不同的时期与上述同行们一起在无线领域进行过研究。在讲授本课程中提出和解决的问题以及在本领域进行的研究,都对本书的写作有重大价值。

作者还要进一步感谢 Sanghyo Kim,是他帮助为本书准备了插图,Qualcomm 的 Edward Tiedemann 帮助组织了 CDMA2000 的内容,IBM 的 Chatschik Bisdikian 博士帮助组织了蓝牙的内容。在休假日,Krishan Sabnani 博士与他在贝尔实验室的无线小组提供了热情好客的支持。作者还要感谢曾经的博士研究生,现在在佛罗里达州立大学的 Li-Hsiang Sun 博士,他承担了第 9 章的部分性能计算工作。最后,作者要感谢对本书手稿指出了很多错误的匿名评论者们,他们还对本书的改进提出了建议。作者希望读者能享受本书的阅读和学习,并学到尽可能多的知识,就像作者在撰写本书时的享受和收获一样。

目　录

第1章　介绍与综述

本书对数字移动无线网络提供了教学性的介绍。这个领域太广泛，并且变化太快，以至于没有哪本书能够包括其所有的方面。然而，本书应该可以为感兴趣的读者提供一个牢固的基础，使之能依此探究本书未包括的内容或对已有的内容进行更详细的讨论。本书附录中的参考文献将提供很多信息，读者应当在需要的时候参考它们。有一家期刊涉及了很多本书所参考的最新工作，这将为读者提供最新的研究结果或最新发展的教学性综述。

注意第1行使用的词汇“数字”。在下面的历史回顾部分，我们将看到最早的无线网络使用的是模拟通信。稍后，本书将对这种模拟网络中的一个——现在仍在使用的 AMPS 进行简要的描述。但是，本书的重点是现代数字无线网络。在世界范围内，现在运营着两种类型的数字无线网络。一种是蜂窝网络类，主要承载话音呼叫，但随着手机或其他基于蜂窝的移动终端对相应业务越来越支持，这类网络也越来越多地承载数据和多媒体业务。虽然不同的蜂窝系统是不能互相比较的，但这些数字无线网络现在无处不在，在世界范围内都能够支持手机用户。第2种类型的数字无线网络是局域网络或个域网络类。本书的大部分内容将关注第1种类型，即无线蜂窝网络，但读者将明显看到本书的很多内容对两种网络都同样适用。另外，第12章对局域网和个域网提供了一个全面的介绍，在局域网类型中集中介绍了使用得越来越广泛的 Wi-Fi 网络，而在个域网中将用蓝牙作为例子。

考虑在后续章节中将要详细描述的蜂窝网络。在这些网络中，用户的手机连接到所谓的基站，每个基站覆盖一个地理小区。基站按顺序连接到有线电话网，从而在原则上允许用户的呼叫传送到世界上任何需要的地方。第1个研发的蜂窝网络是模拟的，从下面的历史回顾部分，我们可以看到它们正被现在世界上最通用的数字蜂窝网络所替代。然面，模拟网络仍然在使用，它们在数字蜂窝无法连接时作为备份。在北美，大量竞争而又互不兼容的数字蜂窝网络可能在一定地理范围或区域内造成某个连接不能进行，这时就会出现上述情况。模拟网络常称为第一代蜂窝网络，现在使用的数字网络称为第二代网络。正如上文所说，虽然后面的章节将会使用一些对所有蜂窝网络(包括局域网和个域网)也适用的材料，但本书的重点是数字蜂窝网络。世界上主要有3种运营的第二代网络：GSM，D-AMPS 或 IS-136，IS-95，这将在第6章和第8章中对其进行详细的描述。

近年来，为了从第二代网络向所谓的第三代网络升级，我们已经做了很多工作。第三代网络使用分组交换模式，而第二代网络使用与数字有线电话网相同的电路交换技术。第10章将详细介绍第三代网络，并将详细介绍电路交换和分组交换的区别。

1.1　历史回顾

在对移动无线网络进行详细研究之前，先对本领域的发展历程做一简要介绍。下文的材

料将重点介绍贝尔实验室的工作，包括在美国的发展。更多的详细内容参见 Frenkiel(2002)[①]和 O'Neill(1985)[②]的第 14 章。这部分的结尾将简要地讨论欧洲从 1981 年起在无线移动通信方面的活动。详细内容请参见 Paetsch(1993)[③]。

船与岸之间的通信属于移动电话技术的最早应用之一，实验性的业务开展于 1919 年美国波士顿和巴尔蒂摩之间的沿岸蒸汽船。采用 4.2 MHz 和 8.7 MHz 幅度调制的商用业务开展于 1929 年，这与幅度调制的电台广播获得大众关注大概发生在同一时期。注意到这些频率相应的波长大概是 70 m 和 35 m，因此轮船成为装载这些波长天线的合适的交通工具。同样，轮船也适合于所需无线设备的大小和重量，并能提供需要的功率。警察通信也大约在同时开展起来。1928 年，底特律警察局引进了一种使用小而结实的无线电台的地面移动通信系统。到 1934 年止，美国 194 个市政系统和 58 个警察系统的 5000 辆警车都装备并使用了移动通信。这些早期的移动通信系统使用的频段是 35 MHz。然而人们很快发现，由于有害的传播效应和大的噪声，使得城市中车辆间的通信常常不能满足要求。城市环境的传播效应是一个未知量，贝尔实验室以及其他地方开始了对它的研究。传播测试最初在 1926 年对 40 MHz 频段开展。到 1932 年，对多个不同频段、多种传输路径和距离的测试都开展起来，并考虑了诸如信号反射、折射、衍射等人们注意到的现象所造成的影响(现在各种研究人员仍然在对不同传播环境中，包括室内和室外的不同频段信号进行测试。第 2 章将介绍传播效应和用来表示这些效应的模型)。1935 年，在波士顿对 35 MHz 和 150 MHz 进行了更多的传播测试。这次对多径效应进行了特别的关注。这些测试同时表明使用调频(而不是先前的调幅技术)能使可靠的传输成为可能。这些各种各样的测试以及在以后年代进行的其他测试，都表明传播效应能够以最简单的形式得到理解，即看成是 3 个因素的组合：一个是形式为 $1/d^n$ 的、与距离成逆向相关的平均接收功率变量，其中 n 是一个比通常自由空间因子 2 要大的整数；一个是平均功率的长期统计变量，现在称为阴影或对数正态衰落；一个是短期的、变化迅速而与车辆移动有关的衰落效应。第 2 章将对这 3 种传播效应进行详细的讨论和建模。

第二次世界大战的爆发中断了无线移动系统的商业活动，然而这些活动在战后得到了快速的发展，特别是在高频段。这些高频段的运营使得更多的用户通道成为可能。1946 年，美国联邦通信委员会(FCC)对圣路易斯的第 1 个商用陆地移动电话系统提供了运营许可证。到当年年底，全美有 25 个城市有该系统的运营。这个基本的系统使用 150 MHz 频段的调频传输，信道载频间隔为 120 kHz。到 20 世纪 50 年代，信道载频间隔减少到 60 kHz，但是由于接收机不能充分地区分相邻的信道，相邻的城市只能交替地使用间隔 120 kHz 的信道。系统之间必须间隔 50 mile(1 mile = 1.6093 km)以上。为了提供与移动用户间的无线连接，人们建造了能覆盖 20 mile 到 30 mile 的高塔。这个系统能提供 40 个信道或并发呼叫。FCC 将这些无线信道均匀划分给本地电话公司(Telcos)和新建立的称为“无线公共业务商”(RCC)移动载频。这些早期的移动系统都是手动操作，通过接线员进行连接呼叫。它们提供半双工的传输，即连接的任何时候都只能有一个方向的通信：双方在连接的空中或无线部分使用同一个频率信道，未

① Frenkiel, R. 2002. "A brief history of mobile communications," *IEEE Vehicular Technology Society News*, May, 4 – 7.

② O'Neill, E. F. 1985. *A History of Engineering and Science in the Bell System: Transmission Technology (1925 – 1975)*, ed. E. F. O'Neill, AT&T Bell Laboratories.

③ Paetsch, M. 1993. *Mobile Communications in the US and Europe: Regulation, Technology, and Markets*, Boston, MA, Artech House.

发声的一方必须通过一个"按键说话"的过程获得信道使用。由于有40个可以使用的信道,根据通话的长度不同,该系统在给定的区域内可以容纳800到1000个顾客(显然,当用户通话长度增加时,平均可容纳的客户数必然减少。这个问题将在本书的后续内容中讨论"阻塞率"这个概念时定量进行研究)。当系统变得越来越流行,想获得一个移动电话将经过一个很长的等待名单。这些系统因而变得有点"精英化",它为那些特别需要成为用户的人们建立了优先权,例如医生。

20世纪60年代新的半导体器件的发明,以及由此引起的系统成本和移动电话功率要求的降低,再加上在移动电话中使用更复杂电路的可能,使得一种非常先进的移动电话业务——IMTS得到了发展,IMTS即所谓的足够逻辑化的"先进移动电话业务"。贝尔实验室从1962年到1964年对这种新的业务进行了开发,并于1965年在宾西法尼亚州的哈瑞斯堡进行了外场测试。这种业务采用间隔30 kHz、截短的调频信道,包含自动拨号,采用全双工模式(即两个方向同时发送),交谈双方都有自己的信道频率。移动电话能够自动搜索空闲的信道,即当前并没有分配给用户的信道。然而,这个新系统在给定区域内仍然只能容纳800到1000个用户,长达25 000个的等待用户名单是非常普遍的。另外,这个业务有限地使用频谱造成了用户常常收到"系统繁忙"的信号,即在分到信道时被阻塞,于是,上文所说的"阻塞率"非常高。

早在1947年,为了缓解当时移动系统容量上的不足,贝尔实验室的工程师们提出了两种解决方案:一种方案是将移动系统转移到更高的频带,提供更多的系统带宽,从而得到更多可用的用户信道;另一种方案是采用地理蜂窝结构。蜂窝系统的概念非常简单,然而其影响是深远的。在蜂窝系统中,一个给定的区域被分成称为小区的邻接地理区域,而信道的频率集合在小区之间被划分。由于信道在足够远的小区之间复用,从而被分配了同样频率的小区间的干扰是可管理的(将在第3章定量研究这些概念)。例如,在第3章所研究的一种方案中,每个小区将被分配总信道数目的1/7。表面上看来好像走到了错误的方向——减少每个小区的信道数目似乎会减少可容纳的并发用户的数目,然而如果每个小区覆盖的面积足够小,特别是在城市地区,则频率能在足够短的距离内复用,这样便能补偿每个小区信道数的减少。但是,我们将在后续章节中考虑一个由此带来的问题:当用户从一个小区漫游到另一个小区时,他们正在进行的通话将要在每个小区分配不同的频率信道。这个信道重新分配的过程称为"切换/handoff"(另一个等价的术语是"切换/handover",它在欧洲较为流行)。对于通话中的用户来说,这个过程必须看起来是无缝的。这样的信道重新分配也要与同一个小区内那些企图获得信道的新呼叫竞争,因而可能会造成正在通话中的用户通话的中断,除非有特别的手段避免或者降低这种情况的几率。本书将在第9章定量讨论切换中断概率并提出了几种保持通话的方式,还比较了它与新呼叫呼损率之间的权衡折中。

让我们回到简短的历史回顾中来。1949年,贝尔实验室向FCC请求允许将移动电话的频带移到470~890 MHz,以获得更多的信道,进而得到更大的移动容量。然而在当时,这个频带是为电视准备的,因此请求被拒绝了。1958年,贝尔实验室请求把764~840 MHz的频段作为移动通信用,但FCC仍然拒绝施行。在此之前,移动系统中引入蜂窝概念已经在贝尔实验室展开了全面的讨论。到1968年止,FCC决定将840 MHz附近的频谱分配给移动电话业务,并为此开启了现代著名的"查询通知和规则制定提出通知"18262号提要。贝尔实验室在1971年做出了回应,提出了一个"大容量移动电话系统"的方案,该方案包含了蜂窝技术的引入(这个方案后来发展成为AMPS高级移动电话业务,即前文所提到的第一代模拟蜂窝系统),于是在

对贝尔实验室提出的这种更高频段的移动业务感到恐慌的不同党派之间展开了10年的冲突。例如，广播公司希望保留此频率分配给广播业务；通信制造商为新系统的前景和可能伴随的竞争而恐慌；无线公共业务商们觉得被贝尔实验室占据了绝对优势；舰队的操作人员希望这些频谱能被他们私人的移动通信使用。直到1981年这些问题才得到解决，FCC最终同意从800～900 MHz的频段分配出50 MHz给移动通信话务使用(实际上最初只分配了40 MHz，其余的10 MHz是在几年后分配的)。这些频率的一半将分配给本地电话公司，另一半分配给无线公共业务商。同时，到当时为止，广泛使用的固态器件、微处理器以及电子电话交换系统，使得车辆定位和小区切换能够在相对小的小区内实行。

在这个"小冲突和政治活动"(O'Neill,1985)的时间里，贝尔实验室关于城市传播效应和城市环境中基于蜂窝的移动系统的工作仍在进行。1978年，在芝加哥对研究所得的AMPS系统开始了技术测试，并在1981年FCC裁决支持系统的发展后，于1983年在同一个城市进行了AMPS的第1次商业铺设(但应该注意，1984年AT&T与其脱离了关系。根据美国司法部和涉及此案的联邦地方法院以及AT&T同意的最终裁决修改案，贝尔系统的移动运营被转移到了RBOC，即当时建立的7个地方性贝尔运营公司)。

如上所述，AMPS系统通常称为第一代蜂窝移动系统，并仍在使用，以作为蜂窝系统的备份。它现在包含两个25 MHz宽度的频带：上行链路或者反向信道(从移动单元到基站)为824～849 MHz，下行链路或者前向信道(从基站到移动单元)为869～894 MHz。如前所述，系统使用模拟调频传输，为每一个方向上的每个信道(即用户连接)分配30 kHz。每个信道最大的频偏是12 kHz。这种通信系统在传输的每个方向上将全部25 MHz带宽划分成30 kHz的信道，称为频分复用系统(FDMA)。频分复用系统以及时分复用(TDMA)和码分复用(CDMA)将在第6章详细讨论，这意味着每个方向上有832个30 kHz的频率信道。根据前文所提到的FCC的裁决，在美国最初铺设的AMPS系统，一半的信道分配给RBOC，即地区性贝尔运营公司，另一半分配给RCC无线公共业务商。同样，如前所述，每个给定的小区分配1/7的信道。例如，考虑一个10 mile半径的小区系统，覆盖大约300 $mile^2$的面积(最初，在芝加哥的实验小区半径为8 mile)。25个邻接的小区将覆盖7500 $mile^2$。而考虑一个作为比较的覆盖相同面积的非蜂窝式的系统，半径大概是50 mile。稍微想想就能明白，在这个例子里小区的引入增大了25/7或3.6倍的系统容量，即可能并发的用户连接或呼叫的数量。使用更小的小区将使系统容量变得更大。

尽管在1983年基于小区的AMPS系统使得移动容量的增加成为可能，但从1985年起，美国的主要城市(比如纽约和洛杉矶)开始面临容量问题。蜂窝电信工业协会开始评价各种替代方案，从1985年到1988年一直都在研究这个问题，最后决定转到数字时分复用系统(TDMA)。如同AMPS一样，这种系统同样在800～900 MHz范围内使用30 kHz的信道，这样是为了向后兼容。然而，每个信道能同时被3个用户使用，从而相应地增大系统容量到以前的3倍(TDMA在第6章讨论)。这种系统的标准得到了发展，并制定了标号为IS-54的标准系统。后来，对标准进行了修改，系统被更名为IS-136，它常被称为D-AMPS，即数字AMPS。这组成了下文将要讨论的3种第二代蜂窝系统之一。第二代数字系统在1991年末开始在美国主要城市运营。对IS-135的详细讨论将在第6章和第8章进行。

1986年，圣地亚哥一个一直在开发码分多址CDMA移动系统的通信公司QUALCOMM得到了很多像NYNEX和太平洋贝尔这样的地区性贝尔运营公司的支持，以试验它们的系统。这

些系统根据 FCC 的规定,必须与 AMPS 和 D-AMPS 一样覆盖相同的 800～900 MHz 的频率范围。这种标号为 IS-95 的 CDMA 系统,从 1993 年开始在美国和其他国家商业化。这种系统是本书将要讨论的第 2 种第二代系统。它在 2 GHz 的 PCS 频段同样被采用(我们将不讨论 2 GHz 频段,为了集中关注于 800～900 MHz 的频段)。在 IS-95 中,分配给移动业务的 25 MHz 带宽被划分成 1.25 MHz 宽度的 CDMA 信道。码分多址将在第 6 章进行详细讨论,并在第 10 章与第三代蜂窝系统联系起来。第 6 章和第 8 章将详细介绍 IS-95 CDMA 系统。

到 20 世纪 90 年代中后期为止,在美国有两个互相竞争并且互不兼容的第二代数字无线系统。如果将后来采用的 2 GHz PCS 频段包括进来,当移动用户在离家很远并且移动载频覆盖很差的地区漫游时,则保持连接将有很多潜在的困难。双模手机可以缓解这个问题:它们能够在某一区域发生通信困难时退回到模拟的 AMPS 系统。这种情况最初在欧洲更糟糕。最早的移动蜂窝系统于 1981 年和 1982 年早期被斯堪的纳维亚半岛国家引入。西班牙、奥地利、英国、荷兰、德国、意大利和法国在 1982 年到 1985 年也有了自己的系统。这些系统是模拟的,但问题是有 8 个这样不相同且不兼容的系统,这意味着通常通信只能限制在一个国家内。这个问题当然很早就被认识到了。1981 年,法国和德国开始了一个发展第二代数字系统的研究。1982 年欧洲邮政电信管理会(CEPT)的电信委员会建立了一个移动通信特别研究小组(GSM),在 900 MHz 频段开发全欧洲范围的第二代数字蜂窝系统的标准。到 1986 年决定使用 TDMA 技术。到 1987 年,欧洲经济共同体采用了最初提出的建议和频率分配,从移动台到基站的上行链路覆盖了 890～915 MHz 的 25 MHz 频段,而从基站到移动台的下行链路覆盖了 935～960 MHz的频段。到 1990 年,所形成的称为 GSM(移动通信全球系统或移动通信 GSM 系统)系统的第 1 阶段标准确定了。另外,同年应英国的请求,将 GSM 适配到 1.8 GHz 频段(DCS1800)的工作开始了。一年后,DCS1800 的标准确定。在 1991 年之前,第 1 个 GSM 系统开始运行,并在 1992 年开始商业运营。GSM 自此开始在整个欧洲配置,允许国家之间的平稳漫游。为北美频段设计的 GSM 版本也投入了使用。当然,这把美国的不兼容问题弄得更复杂,使得 800～900 MHz 频段出现了 3 个不同的第二代系统,并在 2 GHz PCS 频段也有了系统。最初采用 IS-136 的加拿大,也引入了 IS-95 和 GSM。

日本的发展过程和美国相似。当时日本的国有电信公司 NTT 早在 1979 年就引入了模拟蜂窝系统。该系统在 800 MHz 频段的两个 25 MHz 频带内支撑了 600 个 25 kHz 调频双工信道。日本在 1993 年引入太平洋数字蜂窝(PDC)系统,其系统特性近似于 D-AMPS。它与第 1 个模拟系统一样覆盖相同的频段,也在 1.5 GHz 有 1 个 PCS 版本。IS-95 系统也被引入日本,改名为 CDMAOne。

到 2003 年 5 月,GSM 成为世界上使用最广泛的第二代系统,大约有 8.64 亿用户,为世界上 72%的数字移动电话用户(GSM 世界)提供服务。在这些用户中,欧洲有 4 亿,亚太地区有 3.34 亿,非洲和阿拉伯世界各有 2800 万,北美和俄罗斯各有 2200 万。IS-95 CDMA 在世界上排第 2,有 1.57 亿用户使用这种系统。D-AMPS 拥有 1.11 亿用户,日本的 PDC 有 6200 万用户。

1.2　本书概述

结束简短的历史回顾后,我们将为各章节做一个概述。如上所述,无线介质或者说用来实现通信的空间传播条件在无线移动系统的运营和性能中扮演了极其重要的角色。许多年来,

从乡村到城郊，从城市的室外到不同的室内，对在这些不同环境中的传播条件进行了大量的研究，并且这些研究现在仍在进行。这些研究导致了各种模拟不同环境的模型的出现，它们可以在系统设计、实现、仿真和分析中使用。第2章将介绍其中最简单的模型。然而，这些讨论应当有助于理解更加复杂的模型，也应有助于理解关于传播效应方面的最新文献。应尤其关注在给定基站发送信号功率的情况下，统计变化的移动接收机接收信号功率的模型。这个模型以乘积形式综合了3个因素，这些因素概括了无线电波在空气介质中传播时经受的最重要的效应。这3个因素在前面都曾简短地提到。模型中第1项包括了随与发射机距离变化而变化的平均信号功率。与自由空间传播功率的平方成反比不同，发送信号在一个典型的传播环境里与距离的更高次方成反比。这是由于从基站到移动台沿信号传播路径遇到阻碍物而产生的效应，包括反射、衍射以及散射，从而造成了接收机处有多个信号到达。第2章展示的一个例子在后面的章节中也会用到，其中功率随距离的4次方成反比，这是由于从基站而来的直射信号和地面反射的信号叠加形成的效应。

第2章讨论的模型的第2项包括了信号功率在其均值基础上大尺度的变化，有时超过均值，有时落在均值以下。这种效应覆盖了很多波长，并且被发现能够使用对数正态分布对其很好地建模，通常称为阴影或对数正态衰落。这个模型的第3项被设计用来描述小尺度衰落的效果，即当信号的功率随着移动接收机进行波长量级的移动时统计变化。这种衰落是由于发射信号的多径散射造成的，称为瑞利/莱斯衰落，因而接下来的部分讨论信道的特征化。同样，我们将在第2章讨论衰落速度和它与移动速度的联系，以及衰落对携带信息的信号的影响，包括频率选择性衰落条件。这个条件在延时扩展，即不同多径信号的差分延迟超过一个数据符号的持续时间时产生。它会导致信号畸变，特别是信号间的干扰。第2章的结束部分将讨论3种减轻多径衰落效应的方法：为克服符号间干扰的均衡技术、分集过程，以及在CDMA系统中有效地用来改进性能的RAKE接收机技术。

第3章集中关注蜂窝概念以及通过信道重用产生的系统容量的改善。它将信号与干扰之间的比率SIR作为性能参数引入，这在无线系统中通常用来衡量干扰发射机对期望信号的接收所产生的干扰。例如，在蜂窝系统中，如何选择可接受的信干比阈值将决定重用距离，即用同频的小区应当相距多远。当这一章讨论二维系统时，我们将关注六边形的蜂窝结构。六边形通常用来代表蜂窝系统中的小区，因为它们能镶嵌满整个空间，并能近似于使用全向天线时得到的等功率圆。当小区再用距离确定时，对于一个给定的系统，立刻就能得到每个小区的信道数目。通过这个计算，使用阻塞率表示的系统性能就能得到。这个计算决定于每个小区的用户数，以及他们产生呼叫的统计特性。给定一个需要的阻塞率，就能很快得到每个小区允许的用户数，或者需要的小区大小。为了完成这些计算，引入电话系统中关于阻塞率的统计形式——厄兰分布（对厄兰分布的实际推导将留到对性能问题深入分析的第9章）。第3章这个关于性能的介绍性讨论使用了信干比的概念，关注平均信号功率，并忽略了衰落的影响。因此，当这一章结束时，将简要引入概率信号的计算，以确定接收信号的功率超过某个阈值的概率。这些计算将包括第2章描述的阴影衰落模型。

第4章将讨论改进系统性能的其他方法，包括为减小呼损率的动态信道分配策略（DCA）和减少干扰的功率控制。在蜂窝系统中，广泛应用了功率控制来保证维持合适的信干比。从后面的第6章、第8章和第10章对CDMA系统的讨论中看到，功率控制对于其合适的性能是至关重要的。在描述DCA时，我们将集中讨论一个特定的算法，它能够很好地进行近似分析，

却又具有很多 DCA 策略的特性，因此能很好地展示 DCA 是如何改善蜂窝系统的性能的。事实上，我们将把其在简单系统中的使用与在第 3 章对蜂窝信道分配的讨论中暗含的假设，即固定信道分配进行比较。对功率控制的讨论，集中在两种简单的迭代算法，并展示算法的选择是如何戏剧性地影响收敛速度的。然后，将说明这些算法是如何被一种统一的形式写出来的，并与 CDMA 系统中所使用的简单单比特控制算法进行比较。

第 5 章将继续讨论在数字无线移动系统中遇到的基本系统概念，集中讨论这些系统使用的调制技术。本章将从最简单形式的数字调制的简介开始，即相移键控(PSK)、频移键控(FSK)，以及幅移键控或者通断键控传输(ASK 或 OOK)。从这些简单的数字调制技术延伸到正交幅度调制(QAM)技术，我们将第三代蜂窝系统中使用的 QPSK 和 8-PSK 作为特例，同样也将简要介绍数字通信中的波形成形。这些材料对于在数字通信系统方面有些背景的人来说都应该很熟悉。对这些技术进一步的扩展，比如无线系统中使用的数字调制技术 DPSK 和 GMSK 就变得很容易描述了。这一章的末尾将介绍正交频分复用(OFDM)，包括使用快速傅里叶变换技术来实现。OFDM 技术也在高速无线局域网中得到了应用，这在第 12 章中有所描述。

第 6 章将描述数字无线系统中两种主要的多址技术——时分复用(TDMA)和码分复用(CDMA)。TDMA 系统包括重复的时隙帧结构，每个用户在一帧中分配一个或多个时隙。这与现代(有线)数字电话网中所使用的同种技术很相似。GSM 和 D-AMPS(IS-136)蜂窝系统都是 TDMA 系统的范例。以第二代的 IS-95 系统、第三代的 cdma2000 和 WCDMA 系统为范例的 CDMA 系统，使用伪随机码的传输来保证用户的接入。所有系统也都使用了 FDMA 接入，对分配的频段采用指定的频率分配并进一步划分给 TDMA 传输的时隙或在 CDMA 传输中承载多个码。首先对 TDMA 进行介绍性讨论，然后介绍基本的 CDMA(将在第 8 章和第 10 章进一步讨论 CDMA)，接着对 CDMA 提供的潜在系统容量进行简单的计算。这些计算依赖于比特差错率的计算。对不太熟悉通信理论的读者，总结了从噪声中对二进制信号进行检测所得到的经典结论，以及衰落对信号可检测性的影响。此外，我们也简单地描述了第 2 章中讨论的分集技术对差错率的改善。

在这一章的结尾，采用简单的分析模型对 CDMA 的容量进行了计算，使我们能够将 TDMA 和 CDMA 系统的容量性能做出比较。必须强调的是，所得的这些容量结果是假设在蜂窝系统的理想模型下得到的，在实际环境中可能大为不同。然而，在这一章和接下来的其他章节中所使用的系统模型，确实能使我们在铺设无线系统时关注最重要的参数和设计的选择。

为了使读者了解数字无线系统的运行和性能，第 7 章提供了必要的背景介绍材料。这一章很大一部分涉及到了在通信系统和理论的介绍性课程中经常研究的材料。已经掌握了编码理论知识的读者因而可以将这一章的讨论作为复习，并关注这些编码技术在第 10 章所讨论的第三代蜂窝系统中的应用。首先，讨论分组码的纠错和检错，特别关注在无线系统中常用的所谓循环码。接下来是卷积码的讨论，重点在维特比算法和使用卷积编码可能带来的性能改善。第 7 章的结尾将简单讨论第三代 CDMA 蜂窝系统采用的 Turbo 编码。

第 8 章以本书描述的特定数字无线系统的详细讨论开始。本章重点是前面已经多次提到的 3 种第二代蜂窝系统：GSM，D-AMPS 或 IS-136，IS-95。对其中每一种系统，描述移动台和基站之间通过空中接口传输的双向的各种控制信号，这些控制信号用来进行移动台注册和建立呼叫。承载着所需信息的控制信号和业务信号，从为每个类别定义的“信道”中发出。控制信号包括：同步信号、寻呼信号和接入信号。其中，寻呼信号要求某个移动台对进来的呼叫进行

应答，而移动台使用接入信号来请求获取信息传输业务的信道，基站也使用接入信号对此请求进行应答。

对基于 TDMA 的 GSM 和 IS-136 系统，所定义的各种不同信道与在每帧时隙内指定的比特序列相对应。对于每一个系统，不仅会介绍其重复帧结构，也会描述一帧中每个时隙内的比特分配。同时，我们也将介绍不同控制信道传输的各种消息，重点关注建立呼叫的过程。对基于 CDMA 的 IS-95，这些不同的信道与特定的码对应。首先，提供系统业务信道部分的框图描述，从而展开对 IS-95 的讨论。业务信道用于传输实际呼叫信息。然后，转到系统的控制信道，这些信道被移动台用来从基站获取必需的时序信息，从而回答呼叫和请求接入系统。如前所述，对各种基于码的信道及其系统子结构框图的讨论，将进一步深化从第 6 章开始的关于 CDMA 系统的讨论，这样便能通过学习 IS-95 系统来熟悉 CDMA 的概念。同样，如前所述，这些关于 CDMA 的基本内容将在第 10 章讨论第三代系统时得到进一步加强。在讨论完 IS-95 的各种信道框图之后，与如何在 TDMA 中建立呼叫的讨论一样，将展示移动台如何利用各种信道获取必需的相位和时序信息，来向基站注册并建立一个呼叫。不同信道上传播的各种消息的格式也将得到描述。

迄今为止，对第 8 章材料的讨论集中在通过移动台和基站之间的空中或无线接口传播的信号。前述章节的很多材料都着重于无线系统的射频部分，然而这些系统中用户的移动性导致必须在无线系统通常所连接的有线网络上传输信令。在前面提到当移动台跨越小区边界时必须控制所发生的切换，这已经暗示了上述的可能。移动管理在无线蜂窝系统的运行中扮演着重要的角色，不仅包括切换的正确控制，也包含着移动台漫游时的定位，以及定位后对特定小区内的移动台进行正确的寻呼。移动管理的这 3 个方面，需要通过有线网络以及移动台与基站之间的空中接口进行信令传输，这些将在第 8 章中进行详细的讨论。首先讨论的是小区间切换和系统间切换的切换控制，然后转到对位置管理和寻呼的描述，最后通过简单的考虑表明最后两个功能的实现需要在性能上做必要的折中。

第二代蜂窝系统主要用于传输语音电话，尽管数据传输的应用也日渐增多。于是为了在恶劣传输环境中有效且高效地传输话音信号，在第 8 章的结束部分我们将详细讨论所需语音信号的处理。注意，与有线电话传输相比，无线话音信号传输的速率要低得多，这就需要在发射之前先对信号进行大幅度的压缩，然后再编码。所有的第二代系统都采用了各种版本的称为线性预测编码的语音压缩技术。我们将在这一章中简要地描述这种技术，然后继续讨论分别被 GSM，IS-136 和 IS-95 系统采纳的该技术的变化版本。

第 9 章将对前面章节提出的性能问题进行详细的讨论，其背景将是定量描述无线网络中的准入和切换控制，以及在二者之间进行裁决所带来的折中问题。首先从优先原则开始，初步讨论蜂窝中的信道保持时间，并对呼叫持续时间和小区中移动台驻留时间通过一个简单的概率模型来说明，对从宏蜂窝到微蜂窝的各种蜂窝尺寸也都给出了例子。一个小区的平均切换业务量将与采用了流连续性这一参数的新业务的到达联系起来。然后，再基于泊松统计假设和指数概率分布得到新呼叫阻塞率和切换终端率的方程。在分析中顺便导出了前几章提出的厄兰分布。这些基于流连续性参数的分析不涉及特定的小区模型几何形状。接着，使用各种用户端移动性模型对一维和二维蜂窝几何结构的系统性能进行了详细的讨论。我们将给出如何得到重要的统计分布，比如新呼叫和切换呼叫的小区驻留时间分布，以及由此得到的信道持

PRMA。后来提出了许多 PRMA 的改进版本,比如 PRMA++,本章讨论了其中一些接入方案。PRMA 和相关的接入策略都是用于基于 TDMA 的系统的。从前述的第 10 章的概要可以明显看出,蜂窝系统的重点看起来正在向基于 CDMA 的系统转移,因此,在多址接入控制的讨论的最后,我们将介绍特别针对 CDMA 系统而提出的接入策略,其中大多数都来自基本的 PRMA 方案,并在被改进后用于 CDMA 环境。

第 11 章的末尾详细讨论了无线系统中上行(从每个移动台到相应小区的基站)和下行(从基站到恰当的移动终端)的分组数据流的调度过程。调度的目的是两个层面的:在传播条件允许时向用户提供 QoS 性能保证;保证系统资源,主要是链接带宽或容量的完全利用。为传输多个流而进行分组调度的概念在有线分组网络中已经得到长期的研究和实践。在基于分组的无线系统中采用的或提出的一些调度算法,实际上是最初研究的那些用于有线网络的算法的变化版本。本章大量描述并比较了这些算法,而且引用了文献中出现的性能仿真结果。

本书的最后一章(第 12 章)将详细讨论承载高比特率数据业务的无线局域网(WLAN)和以蓝牙标准为范例的无线个域网(WPAN)。无线局域网提供了大约 100 m 的覆盖范围。个域网用于距离最多 10 m 的不同设备间的无线通信。本章对 WLAN 的讨论将关注 IEEE 802.11 标准,重点在高度成功和广泛使用的 802.11b,通常称为 Wi-Fi。802.11b 标准设计运作的标称数据为 11 Mbps。允许数据传输率高达 54 Mbps 的更高比特率的 WLAN 标准,802.11g 和 802.11a 也将被讨论。这些更高比特率的 WLAN 标准使用了第 5 章描述过的 OFDM 技术。蓝牙标准也被 IEEE 采用成为 IEEE 802.15.1 WPAN 标准,设计运作于 1 Mbps 的速率。两个标准用的都是 2.4 GHz 免许可证的无线通信频段。

第 12 章的第 1 部分讨论 IEEE 802.11 无线局域网标准。它首先简要介绍有线局域网通用的接入技术——以太网,然后对 802.11 所采用的策略——载波侦听多址接入/冲突避免(或称 CSMA/CA)进行了详细的描述。冲突避免机制是必要的,因为接入同一个无线信道的多个用户会产生冲突。以太网和 802.11 的接入策略都是操作在基于分组分层架构的介质接入子层(MAC)的,这种架构就是第 10 章介绍的架构。实际数据传输发生在这种架构的物理层,关于 802.11b,802.11g 和 802.11a 的物理层规范的描述,被安排在介绍 MAC 层接入控制的段落之后。其中 802.11g 和 802.11a 两个标准是基于 OFDM 技术的。在给出 802.11 标准的描述之后,我们将继续给出该标准的数据吞吐量性能分析,并用一些例子来阐明这些讨论。

第 12 章的第 2 部分和结论部分详细地描述了蓝牙和 IEEE 802.15.1 无线个域网标准。该标准规定蓝牙器件在 10 m 的覆盖范围内将通过自组织方式形成"微微网"[①],一个微微网包括一个主控台和最多 7 个从控台,每个从控台与主控台通信,并在主控台的控制下运行,因此在这个系统中没有竞争。这一节的讨论包括微微网的建立方法,用于建立微微网的控制分组的描述,以及在微微网建立之后进行通信的各种数据业务分组的格式。本节还包括关于蓝牙性能的讨论,这些讨论建立在文献中对蓝牙性能的仿真研究基础之上。很多仿真结果包括了阴影衰落和莱斯类型的快衰落的影响。第 2 章进行过关于各种类型的衰落的讨论。在数据分组传输的微微网中,主控台的调度对蓝牙的性能起到了决定性作用,因此,文献中出现的多种调度技术之间的比较也在本节一并给出。

---

① 也称为"皮网"——编者注。

第2章　无线介质中传播现象的特性

前一章简要介绍了本书中将要讨论的各章内容,可以看到无线传播环境在无线移动通信系统中起着重要的作用,它决定了系统的性能,决定了系统是否适合进行实时语音信息传输,或者适合数据传输,或者其他类型的通信。所以,在介绍无线通信其他方面之前,有必要对无线介质某些方面的影响详细地讲解一下。在第1章中,可以看到无线通信或无线线路一般都被描述为无线系统中移动用户终端和它所通信的基站之间的通信链路,而基站则连接着可传输信号的有线网络。现代的通信系统一般在区域上被分为很多小区,每个小区设有一个基站。在以后的章节中我们将会详细讲述小区和蜂窝结构(例外的是第12章主要讨论小型无线网络,基本不涉及基站和蜂窝的概念)。本章主要讨论一个小区内信号在基站和移动终端之间进行传输时的传播环境。

无线链路由收发双用信道组成:前向通路即下行,是从基站到移动终端;反向通路即上行,是从移动终端到基站。由于电磁波在传播过程中经常会遇到障碍发生反射、衍射、散射等,这使得不管是上行链路还是下行链路,其信号传播环境都很难刻画出来,这样得到的接收信号往往变化得很随机。当终端移动时,信道环境也发生了变化,使得信号电平随机波动,这种现象就是信号衰落。衰落速率与移动终端对基站的相对移动速度相关,也与所传输信号的频率相关。在用不同的传播模型来近似模拟不同的物理环境方面,前人已经做了很多工作。这些模型可以用来决定在一个给定区域内,使用多少基站,怎么安排基站位置,才能使移动用户得到最好的服务,已经有基于这些模型的软件可以完成这样的任务。另外,它们也可用于对已搭建的无线系统的性能进行评估。

在本书中,为了便于大家对移动通信的过程有一个基本的认识和理解,我们主要介绍最简单的模型。本章的第1节将主要回顾自由空间的信号传播,接下来讲述的是本书中使用的传播模型,在无线通信中它有3个一般特征:平均功率方差是 $1/(距离)^n$,其中 $n>2$,而 $n=2$ 为自由空间的情况;平均功率的长期变化或衰落,称为阴影或对数正态衰落;短期多径衰落导致瑞利/莱斯衰落,其功率变化发生在一个波长范围内。瑞利衰落的引入可以是多径无线介质或常说的信道被表示成为一个有线性冲激响应的模型。然后,讨论衰落速率及其与移动终端速度的关系,还有衰落对承载信息的传输信号的影响。接下来,特别讨论了频率选择性衰落的情况,以及与之相关的接收信号的时间色散和时延扩展。对于数字信号,频率选择性衰落将会导致符号间的干扰。在这节的结尾,大致总结了一些减轻信号衰落的方法,包括信道均衡、分集接收以及瑞克时间分集机制。

2.1　自由空间传播情况

众所周知,射线光束可以在一个远场区域有天线的接收系统中用来发现信号的功率变化。特别地,对于一个各向同性的天线,发射功率为 P_T,那么在距离为 d m的地方,接收到的功率谱密度为 $P_T/4\pi d^2$ w/m²。当接收端的有效接收面积为 A_R 时,接收功率 P_R 可以表示为

$$P_R = \frac{P_T}{4\pi d^2} A_R \eta_R \tag{2.1}$$

其中 $\eta_R < 1$,是有效作用系数。图 2.1(a)就描述了这种情况。

运用各向同性发射器,大多数的天线都有发射增益,这可看做是各向同性情况的聚焦效应或聚焦增益。如果一个发射天线的增益 $G_T > 1$,那么接收面积为 A_R 的接收功率为

$$P_R = \frac{P_T G_T}{4\pi d^2} A_R \eta_R \tag{2.2}$$

发射增益 G_T 与发射天线的有效发射面积 A_T 成正比(发射电磁波束所成的立体角与天线有效面积成反比)。当天线的大小以发射或接收电磁波的波长来计量时,其大小决定了天线的聚焦性能,天线越大,即波长数越多,发射的波束越密集,能量越集中。发射增益与面积之间的关系为 $G_T = 4\pi \eta_T A_T / \lambda^2$,其中 $\eta_T < 1$ 是发射天线效率因子。有效接收面积 A_R 与接收增益也有类似的关系,如果式(2.2)中 A_R 由表达式 $G_R = 4\pi \eta_R A_R / \lambda^2$ 重写,那么可得到自由空间接收功率的表达式

$$P_R = P_T G_T G_R \left(\frac{\lambda}{4\pi d}\right)^2 \tag{2.3}$$

例如,这一等式常用于卫星通信中确定所需功率。图 2.1(b)描绘了定向通信的情况。

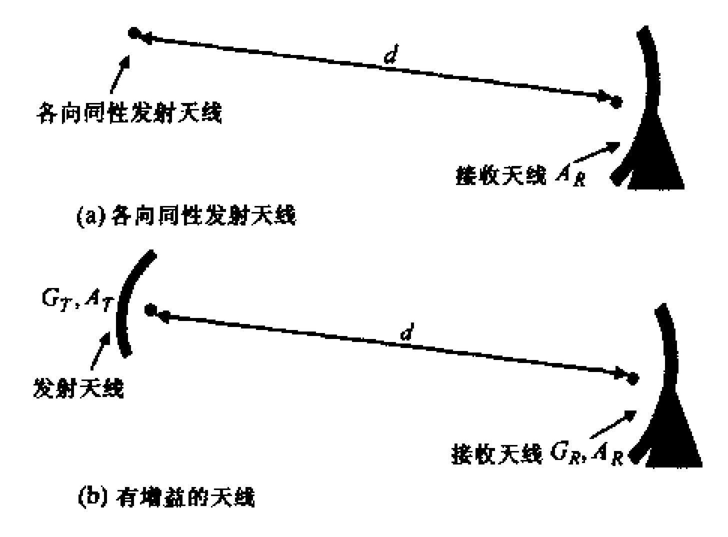


图 2.1 自由空间通信

2.2 无线情况

现在讨论无线情况下的电磁波传播,为简单起见,主要讨论下行链路即从基站到移动终端的情况,上行链路即从移动终端到基站的情况与之类似。前面已经提过,在一个给定小区内,在基站发射信号到达假定的无线移动终端之前,经常会遇到多种障碍物,比如树木、建筑物、车辆等。发射的电磁场遇到这些各样的障碍物后发生反射、衍射或散射,使得接收机接收到的信号为各种干扰信号的叠加。接收机所得到的电磁场是由空间受到混合干扰的各路电磁波组成的,于是,从接收端看传输信道,可以看成是一个传输能量空间变化的模式。当接收端移动时,它所接收到的信号也会变化。当它从相对高功率(优良接收)的区域移动到低功率(劣质接收)

的区域时,就发生了衰落。举一个极端一点的例子,假设一个移动接收机被一个人带着或者是被车载着,正在市内一个拥挤的环境中转弯,这时接收天线收到的电磁波将会有剧烈的变化,接收机正处理的信号特性也会有明显的改变。

移动无线接收机通过这种"严酷"的传播环境时所产生的影响将会相当大地改变自由空间的功率变化等式,有如下一些具体影响:

1. 在距离超过很多个波长的情况下,测得的远场平均功率随着距离的增加而减小,其减少的速率大于 d^2,举一个一般的例子,远场平均功率以 $1/d^4$ 的速率减小。
2. 在相对远的距离同样超过很多个波长的情况下,测得的实际接收功率在平均功率左右随机变化。这时,一个很好的近似就是假设以分贝(dB)为单位的功率,服从以平均功率为均值,以 6 dB 到 10 dB 范围内的某个数为方差的高斯分布或正态分布。那么,功率概率分布就称为对数正态分布。前两个影响远场平均功率和实际接收功率,一般是涉及大尺度衰落在相对远的距离下发生变化。
3. 在很近的距离内,波长范围内可测得的信号将大幅度变化。比如,接收机以 $\lambda/2$ 为步长移动,接收的信号可能相差很多个分贝。对于移动的接收机,信号在小距离内发生的大变化是快速发生的信号变化,信号变化的速率与接收机移动的速率成正比。这种信号的小尺度变化是由多路接收信号结合而成的有害或者有益的相位干扰所造成的,这种现象涉及到多径衰落。对于相对大的小区系统,称为宏小区系统,由于多径而引起的接收信号幅度变化经常用瑞利分布来模拟。而对于很小的小区内的系统,称为微小区系统,用莱斯分布可以很好地近似小范围信号的随机变化。以后将详细介绍这些分布。

将这 3 种现象组合在一起考虑,在无线蜂窝系统中,统计意义上的接收功率可由下面等式确定:

$$P_R = \alpha^2 10^{x/10} g(d) P_T G_T G_R \tag{2.4}$$

其中,$10^{x/10}$ 和 α^2 分别体现了阴影衰落和多径衰落的影响,x 和 α 都是随机变量,这两类衰落将在后面讨论。$g(d)$ 项也将在下面详细讨论,它体现了功率和距离成反比的关系。另外,$g(d)P_TG_TG_R$ 表示平均功率,与距发射机为 d 的接收机收到的功率一样,即平均接收功率可以表示为

$$\overline{P}_R = g(d) P_T G_T G_R \tag{2.5}$$

在自由空间,$g(d)$ 就是 $1/d^2$,而对于一个普通的两径模型,$g(d)=kd^{-4}$,其中 k 是常数。更普遍地,可以有 $g(d)=kd^{-n}$,其中 n 是一个整数,这一表达式常用于计算宏小区系统的性能。对于微小区系统,某些人用下面的表达式来代替 $g(d)$:

$$g(d) = d^{-n_1}\left(1+\frac{d}{d_b}\right)^{-n_2} \tag{2.6}$$

其中,n_1 和 n_2 是两个独立的整数,d_b 是一个已测得的极限,表 2.1 显示了在 3 个不同城市中适合测试的 n_1,n_2 和 d_b 的不同值(Linnartz, 1993)[①]。

① Linnartz, J. P. 1993. *Narrowband Land-Mobile Radio Networks*, Boston, MA, Artech House.

表 2.1 经验功率下降值

城市	n_1	n_2	d_b(m)
伦敦	1.7~2.1	2~7	200~300
墨尔本	1.5~2.5	3~5	150
奥兰多	1.3	3.5	90

另外一个简单模型的 $g(d)$ 表达式如下：

$$\begin{aligned} g(d) &= d^{-n_1} & 0 \leqslant d \leqslant d_b \\ &= d_b^{-n_1}(d/d_b)^{-n_2} & d_b \leqslant d \end{aligned} \tag{2.6a}$$

对于式(2.6a)的情况，以分贝(dB)测量的平均接收功率为

$$\overline{P}_{R,\mathrm{dB}} \equiv 10\lg \overline{P}_R$$

图 2.2 所示的是一个双线性斜率的形式，在距离为 d_b 时斜线发生转折。

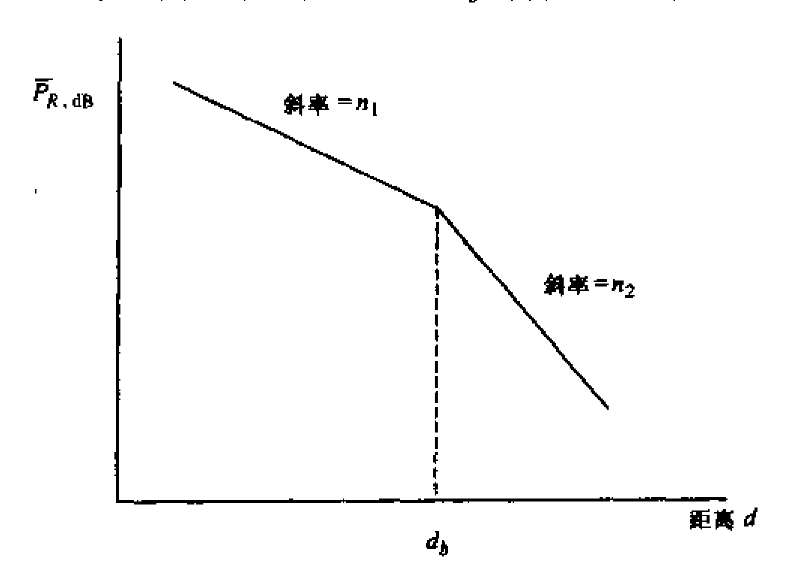


图 2.2 无线通信中双斜率的接收信号模型

从前面已经知道，由式(2.4)得到的 $\overline{P}_R = P_T g(d) G_T G_R$ 提供了一个平均功率(或者等效的功率密度或能量密度)的测量方法。这一功率表达式有时称为面积上的平均功率。前面已经指出，式(2.4)所给的瞬时实际接收功率 P_R 在平均功率左右随机变化。对于大尺度统计变化，阴影或者对数正态衰落而言，由 $10^{x/10}$ 这一项说明；对于小尺度变化或者多径衰落则由 α^2 体现。在下面的段落中我们将分别讨论这两项，先是阴影衰落，然后是长期路径损耗和小尺度衰落。

阴影衰落

先考虑式(2.4)中的阴影衰落项 $10^{x/10}$，前面已经指出，这种衰落变化相对缓慢，在相对远的距离(许多个波长)才能体现出来。由于以分贝为单位的接收信号功率可以由一个高斯随机变量来近似地表示，于是式(2.4)中这一有些特殊的表达式可以从这个角度来进行解释。以分贝为单位的 P_R 表达式如下：

$$\begin{aligned} P_{R,\mathrm{dB}} &= 10\lg P_R \\ &= 10\lg \alpha^2 + x + 10\lg g(d) + P_{T,\mathrm{dB}} + 10\lg G_T G_R \\ &\equiv 10\lg \alpha^2 + p_{\mathrm{dB}} \end{aligned} \tag{2.7}$$

以分贝为单位的阴影衰落随机变量 x，是一个零均值、方差为 σ^2 的高斯随机变量

$$f(x)=\frac{e^{-\frac{x^2}{2\sigma^2}}}{\sqrt{2\pi\sigma^2}} \tag{2.8}$$

下面表达式给出了统计波动意义下的以分贝为单位的平均接收功率或面积上的平均功率

$$\overline{P}_{R,\mathrm{dB}}=P_{T,\mathrm{dB}}+10\lg g(d)+10\lg G_T G_R$$

暂时忽略掉体现多径衰落影响的随机项 α^2，式(2.7)中的 p_{dB}项，则这一随机变量体现了以 dB 分贝为单位的由于阴影衰落而统计变化的长期接收功率，它是均值为 $\overline{P}_{R,\mathrm{dB}}$的高斯随机变量（这里忽略了移动时变化很快的多径衰落，后面会提到。多径衰落是在波长级距离范围内变化的，阴影衰落是以米为单位变化的，这点前面已经提到过，下面的讨论也将见到）。一些时候 p_{dB}项称为局部平均功率，与面积上的平均功率区别开来，同时也说明它与包含变化很快的多径衰落影响的瞬时接收功率 P_R 不同。p_{dB}的概率密度函数如下：

$$f(p_{\mathrm{dB}})=e^{-(p_{\mathrm{dB}}-\overline{P}_{R,\mathrm{dB}})^2/2\sigma^2}/\sqrt{2\pi\sigma^2} \tag{2.9}$$

图 2.3 给出了一个典型的以分贝为单位的功率图，它是以基站和移动接收机间的距离 d 为变量的函数，显示了阴影衰落下接收功率在平均功率附近的统计散布，忽略了小尺度多径衰落。如前所述，σ 的典型值是 6 dB 到 10 dB 范围内的值。图 2.3 显示的是 $g(d)=kd^{-n}$ 的情况，其中 n 是一个整数，所以平均功率与距离关系的曲线如图所示是一条反比直线。

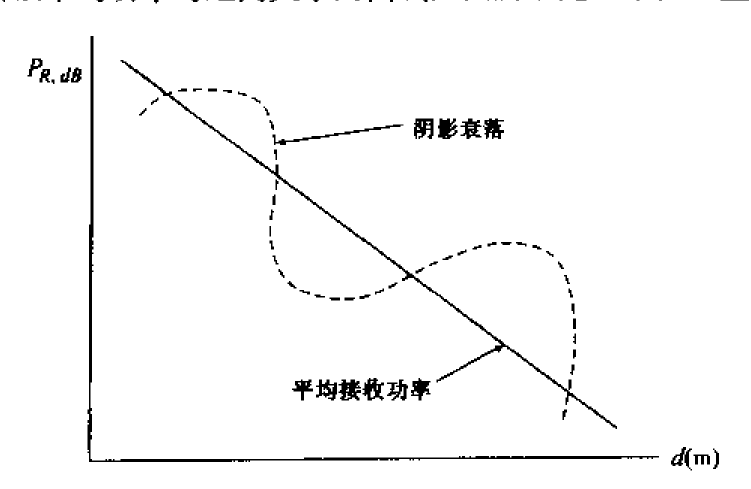


图 2.3 无线传输中的大尺度传播效应

一般情况下，长期阴影衰落是指信号遇到地理障碍物像山峰，或者人工障碍物像建筑，面引起的传输信号的功率变化。所以，即使在与发射机相同距离的不同位置，所测得的信号功率可能相差很大。长期衰落导致了传输信号的域强度在平均功率和由 $P_T g(d) G_T G_R$ 给出的面积上的平均功率的附近波动。可以说，阴影效应显示了几十米到几百米范围内无线通信域强度的中比例波动。许多经验研究，像多年前由 Reudink 和他的合作者们(Black 和 Reudink，1972)[①] 以及 Egli(1957)[②] 所做的一样，显示出信号的接收平均功率在面积上的平均功率周围以对数正

① Black, D. M. and D. O. Reudink. 1972. "Some characteristics of radio propagation at 800 MHz in the Philadelphia area," *IEEE Transactions on Vehicular Technology*, 21 (May), 45 – 51.

② Egli, J. J. 1957. "Radio propagation above 40 Mc/s over irregular terrain." *Proc. IRE*, October, 1383 – 1391.

态分布波动，于是验证了式(2.7)到式(2.9)所给出的模型。这样的研究和经验结果都是基于在距发射机固定距离处所接收到的信号强度的测量，这些测量都是在运载车辆沿圆周或半圆周运动时每百米重复测量的结果。图 2.4 和图 2.5 给出了在距发射机接近 2 km 距离的实验结果(Jakes, 1974)[①]。从这些图中可以看到一个很有意思的现象，由于树木叶子的有或无，夏天和冬天所测得的接收信号强度相差 10 dB。

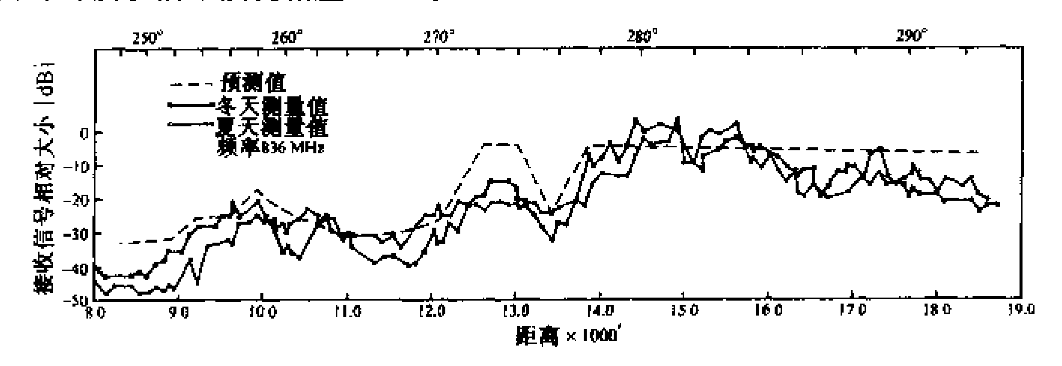


图 2.4　836 MHz，距发射机固定距离下测得的信号强度(引自 Jacks, 1974; Fig.2.2－24)

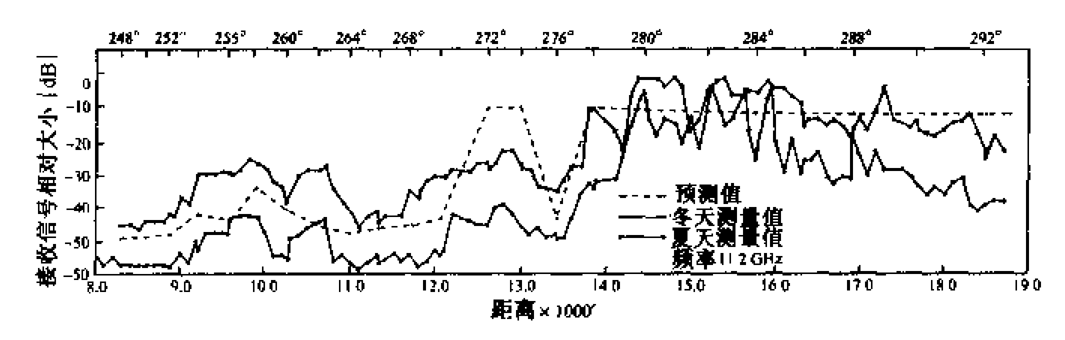


图 2.5　11.2 GHz，距发射机固定距离下测得的信号强度(引自 Jacks, 1974; Fig.2.2－25)

如上所述，阴影效应主要是由地形和建筑引起的，由于每一个地理区域都有其自己独特的地形、建筑物、建筑物的形状和密度，以及建筑材料，所以式(2.8)和式(2.9)高斯模型中的方差 σ 取决于具体的环境。基于这一现象，进行区域实验测量以及在使用蜂窝系统前精心描述环境特征都是很重要的。在本书中，为了简单起见，假设已知值 σ 的范围为 6～10 dB。读者在具体使用时需注意，以后章节中所分析得到的结果一般都是基于简化模型的。

路径损耗：两径模型

现在要更详细地讨论平均功率传播项 $g(d)$，集中讨论普通的两径传播模型，在此模型中 $g(d)=kd^{-4}$。这一模型是文献中最常用于蜂窝系统性能计算的模型，在本书中也将使用这一简单模型进行性能计算。

由于反射、衍射和散射而在接收机得到的多路波束对平均接收功率产生了影响，两径传播模型是可用于展示这种影响的最简单模型，它涉及了一个反射波。尽管这个模型相对简单，但是在相对高的基站天线和/或基站用户之间可视的宏小区系统中，它提供了合理的精确结果。在两径模型中，假设发射电磁波一路是通过可视路径直接到达接收机，另一路是在平坦地面上理想反射后间接到达接收机。图 2.6 描绘了这一几何模型，包括基站的发射天线，离地面高度

① Jakes, W.C. ed. 1974, *Microwave Mobile Communications*, AT&T, 1995 edition, New York, IEEE Press.

为 h_t；地面为理想反射的平坦表面；在自由空间距离为 d m 的地方是高为 h_r 的接收天线。最初为简单考虑，假设两个天线都是各向同性的发射器。后面再考虑方向性天线的适当增益因子，并与自由空间情况进行比较。间接的波束在到达地面之前传播了 d_1 的距离，反射波在到达接收天线前传播了 d_2 的距离。入射角和反射角为 θ，发射天线和接收天线间距离为 r。

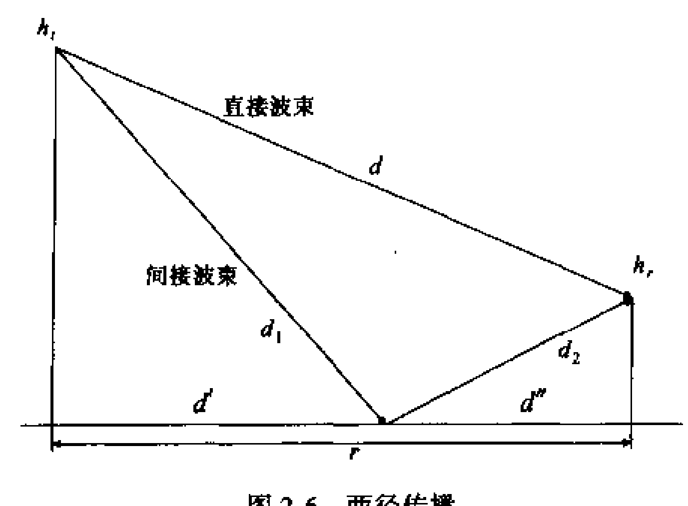


图 2.6　两径传播

假定对于发射机远场中的电场是一个频率为 f_c、幅度为 E_T 的正弦波，其复数符号表达式如下：

$$\tilde{E}_T = E_T e^{j\omega_c t}$$

其中，$\omega_c = 2\pi f_c$。发射功率 P_T 与电场幅度的平方成正比，为 KE_T^2，K 是一个比例常数。现在来看直接到达接收天线的波束，其复相位表达式为

$$\tilde{E}_{R,D} = \frac{E_T}{d} e^{j[\omega_c(t-\frac{d}{c})]} \tag{2.10}$$

其中，c 为光速。对于接收到的间接波，由于假设在地面发生理想反射，其形式比较简单，除了它传播总长为 $d_1 + d_2$，同时由于理想反射，它的相位增加了 π。于是，其表达式如下：

$$\tilde{E}_{R,I} = -\frac{E_T}{d_1 + d_2} e^{j\left[\omega_c\left(t-\frac{d_1+d_2}{c}\right)\right]} \tag{2.11}$$

总的接收电磁场是这两个相量的和，其式如下：

$$\tilde{E}_R = \frac{E_T e^{j\omega_c(t-\frac{d}{c})}}{d}\left[1 - \frac{d}{d_1 + d_2} e^{-j\omega_c\left(\frac{d_1+d_2-d}{c}\right)}\right] \tag{2.12}$$

注意到平均接收功率 P_R 与电场大小的平方成正比，可从式(2.12)扩展此结果到带增益的有向天线，平均接收功率表达式如下：

$$\overline{P}_R = K|E_R|^2 = P_T G_T G_R \left(\frac{\lambda}{4\pi d}\right)^2 \left|1 - \left(\frac{d}{d_1 + d_2}\right) e^{-j\Delta\phi}\right|^2 \tag{2.13}$$

这里 $\Delta\phi \equiv (d_1 + d_2 - d)\omega_c/c = 2\pi\Delta d/\lambda$，其中 $\Delta d \equiv (d_1 + d_2 - d)$，电磁波的波长 $\lambda = c/f_c$。具体推导由读者完成。

现在用图 2.6 和一些简单的代数来表明 $d_1 + d_2$ 可由发射机与接收机间距 d，以及它们的高度 h_t 和 h_r 来表示。在合理的假设下，d 相对于天线高度是较大的，则式(2.13)可简化为

$1/d^4$ 的函数关系。首先 $d_1^2 = d'^2 + h_t^2$ 且 $d_2^2 = d''^2 + h_r^2$，距离 d' 和 d'' 如图 2.6 所示，然后由间接波的入射角 θ 与反射角相等，得到 $d' = d''h_t/h_r$。另外，$d' + d''$ 就是发射机和接收机之间的地面距离 r，于是用 r 来替换 d' 和 d''，可得到

$$d_1 + d_2 = \sqrt{r^2 + (h_t + h_r)^2} \tag{2.14}$$

从图 2.6 的几何图形可以清楚地得到 $d^2 = (h_t - h_r)^2 + r^2$，于是可以将式(2.14)中的 r 消去得到

$$d_1 + d_2 = d\sqrt{1 + \frac{4h_t h_r}{d^2}} \tag{2.14a}$$

现在假设 $d^2 \gg 4h_t h_r$。下面验证一下，如果发射天线高 h_t 为 50 m，接收天线高 h_r 为 2 m，那么意味着 $d \gg 20$ m，如果发射天线高度增为 100 m，那么 $d \gg 30$ m，这一结果很合理。这意味着式(2.13)中 $(d_1 + d_2)/d$ 近似为 1。现在考虑式(2.13)中 $\Delta\phi$ 的影响，注意到 $\Delta d = (d_1 + d_2 - d) \approx 2h_t h_r/d$，于是 $\Delta\phi \approx 4\pi h_t h_r/\lambda d$。现在假设这一相位差项很小，特别地，假设 $\Delta\phi \leqslant 0.6$ rad，这就使得 $d \geqslant 21h_t h_r/\lambda$，一般情况下对于 d 会有更严格的限制。特别地，取工作频率 $f_c = 800$ MHz ($\lambda = 3/8$ m)，这一频率大约是大多数第一代和第二代蜂窝系统簇所使用的频率。对于上面的假设 $h_t = 50$ m，$h_r = 2$ m，可得 $d \geqslant 5.6$ km。如果使用更高的工作频率如 2 GHz($\lambda = 0.15$ m)，则 $d \geqslant 14$ km。在很小的 $\Delta\phi$ 和 $(d_1 + d_2)/d = 1$ 的假设下，得到

$$\left|1 - \frac{d}{d_1 + d_2}e^{-j\Delta\phi}\right|^2 \approx (\Delta\phi)^2$$

式(2.13)于是可简化为

$$\overline{P}_R = P_T G_T G_R \left(\frac{\lambda}{4\pi d}\Delta\phi\right)^2 = P_T G_T G_R \frac{(h_t h_r)^2}{d^4} \tag{2.13a}$$

注意到在最后得到的平均接收功率的表达式中存在着 $1/d^4$ 这一与距离相关的函数关系。$(h_t h_r)^2/d^4$ 项即是式(2.4)和式(2.7)中的 $g(d)$ 项。前面已经指出，$1/d^4$ 这一平均功率函数关系经常用于蜂窝系统的性能研究中。在本书接下来的内容中将介绍很多使用了它的例子，不过，还要注意到这里还有另外一个假设

$$r \approx d \approx d_1 + d_2$$

于是，发射系统(基站)和接收系统(移动终端)之间的地面距离 r，可以在计算平均接收功率时替代距离 d。实际上，我们将经常用到系统间的地面距离。

现在考虑一些测量数据来大致地验证一下这些两径结果，图 2.7 给出了在 900 MHz 下使用较高天线时，纽约、费城、东京三个城市以距离为自变量的接收信号功率函数的曲线(Jakes, 1974)。这些测量结果与式(2.1)所表示的自由空间传播模型进行了比较，自由空间传播中功率下降率为 $1/d^2$。图 2.8 给出了在东京 922 MHz 下，基站天线高度(在图中以符号 h_b 表示)对于接收信号强度的影响，也是接收到基站间距离的函数。注意到这些曲线大致符合由两径模型得到的简单表达式(2.13a)，这些曲线还说明传播损耗确实要比自由空间中下降得快，同时也显示出高的基站天线会增加平均接收信号功率，像式(2.13a)中给出的一样。

多径衰落：瑞利/莱斯模型

本节将讨论在无线系统中由于无线电波的散射、衍射和反射面产生的 3 个显著传输影响。

这 3 个影响在简化接收功率表达式(2.4)中都有体现,一个影响就是平均功率与距离的关系 $1/d^n$,只讨论了 $n=4$ 这个一般的情况,这一情况是由多径模型得到的。在讨论对数正态衰落或者阴影衰落之前,常常先给出平均功率的大尺度随机变化。现在来看第 3 种现象,小尺度多径衰落,将使用瑞利/莱斯统计量模型对其进行描述。前面已经指出,其幅度和功率的变化是在波长级距离范围内的,所以涉及小尺度衰落。特别地,对于瑞利统计,随机变量 α 的概率密度函数 $f_\alpha(\alpha)$,$\alpha\geqslant0$,如下所示:

$$f_\alpha(\alpha)=\frac{\alpha}{\sigma_r^2}\mathrm{e}^{-\alpha^2/2\sigma_r^2}\qquad \alpha\geqslant0 \tag{2.15}$$

其中,σ_r^2 是可调整的瑞利参数,可以很容易地看出瑞利分布的二次阶矩为 $E(\alpha^2)=2\sigma_r^2$,瑞利函数的略图如图 2.9 所示。莱斯统计量将在后面讨论。对于式(2.4)给出的接收功率模型,可以看出必须设置 $\sigma_r^2=1/2$,为了证明这点,需要回到基本传播公式式(2.4)中。前面已经提到多次,平均接收功率是 $P_T g(d)G_T G_R$,这意味着将瞬时接收信号功率 P_R 关于随机变量 x 和 α 求期望,然后就可以得出 $E(\alpha^2)=1$,于是需要设 $\sigma_r^2=1/2$。另外一种方法就是考虑以分贝为单位的 x 和 α 的接收功率表达式式(2.7)。可以很清楚地看到 $E(x)=0$ 和 $E(\alpha^2)=0$,这时也可以看到瞬时局部平均接收功率 P_R 是一个呈指数分布的随机变量。

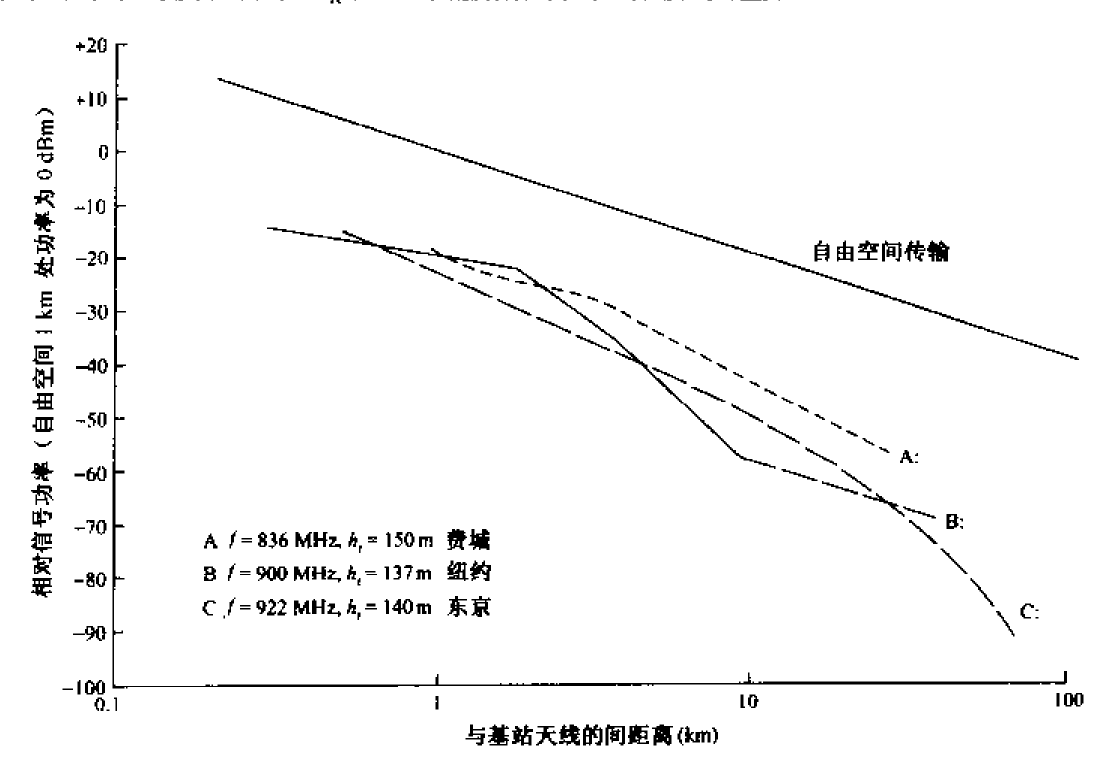


图 2.7　以距离为自变量的接收信号功率(引自 Jacks,1974:Fig.2.2-10)

对于随机多径,很自然地就会产生瑞利分布,在移动电话早期就发现测量结果服从瑞利分布(Jakes,1974)。实际上,由于在给定路径上遇到建筑或者障碍物发生散射,使得每一个波束都是由多路波束叠加而成的。考虑图 2.6 中所示的直接到达接收天线的波束,由于在它传播的路径上遇到建筑或者其他物体发生多次散射,所以它实际上也是由很多波束组成的,而且这些波的平均传播距离为 d,示意图如图 2.10 所示。在接收端接收到的每一路波束,由于散射在幅度和相位上都会产生随机差异。另一方面,可以考虑图 2.3 所示的阴影衰落的例子,随机变化的、在阴影

衰落曲线上的每一个功率点实际上都会由于多路波束散射而在波长级距离范围内随机变化。现在关注的是多波束现象,假定有 L 路这样的波束,如果发射信号还是频率为 f_c Hz 的未调制的正弦波,则由式(2.10)可写出 L 路波束叠加的复相量表达式,在接收端得到的合成相量式为

$$\tilde{S}_R(t)=\sum_{k=1}^{L}a_k e^{\mathrm{j}[\omega_c(t-t_0-\tau_k)+\theta_k]} \tag{2.16}$$

其中,$t_0=d/c$ 项代表了式(2.10)中的平均时延,L 路波束中每一路相对这一平均时延都有一个随机差值 τ_k。假定随机相位 $\omega_c\tau_k$ 与随机附加相位 θ_k 一样均匀分布于 0 到 2π 之间,则这两项的和 $\theta_k-\omega_c\tau_k\equiv\Phi_k$ 也均匀分布于 0 到 2π 之间,第 k 路波束的幅度 a_k 是一个实随机数(很快将会讲到,移动终端的运动将会引入附加的多普勒相移)。

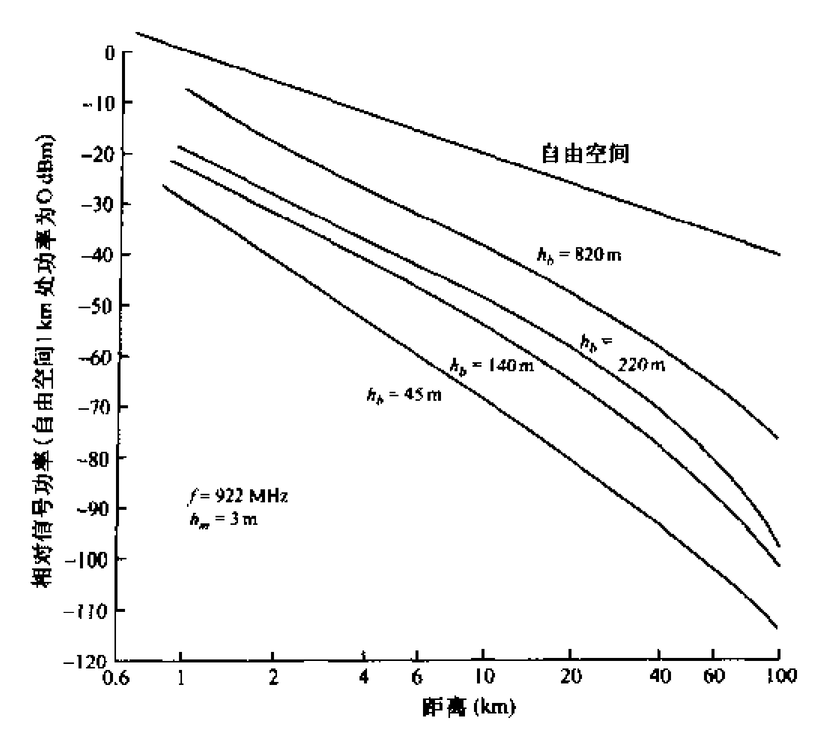


图 2.8 基站天线高度对接收功率的影响(引自 Jacks,1974:Fig.2.2-11)

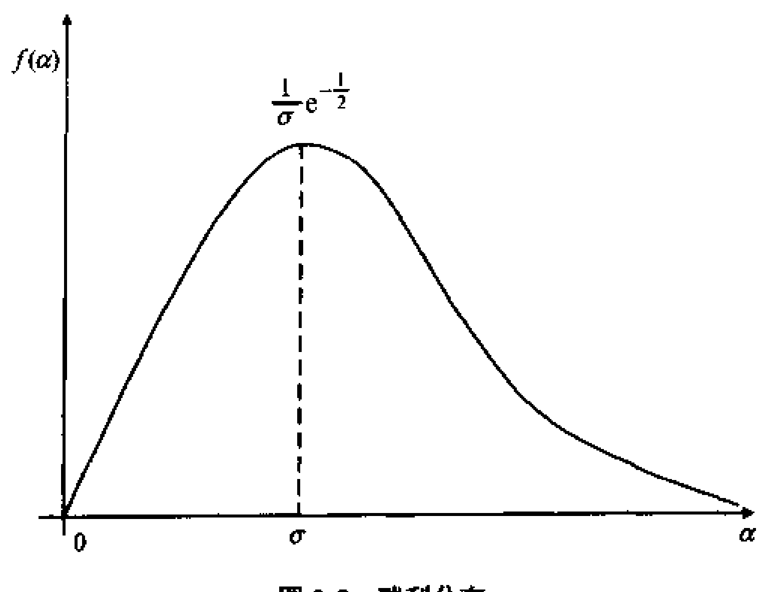


图 2.9 瑞利分布

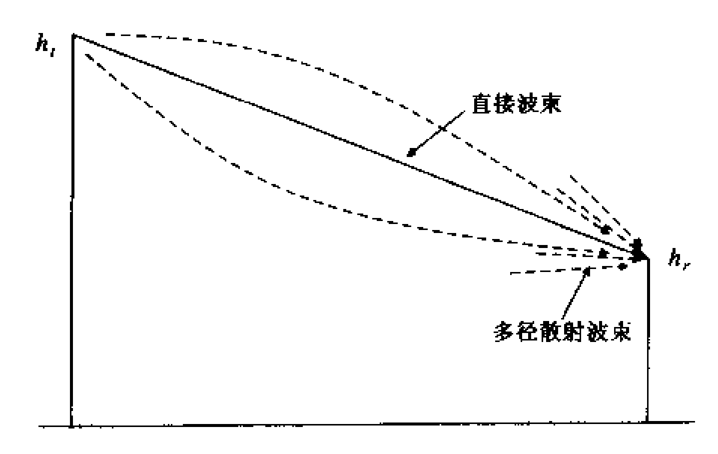


图 2.10　接收到的多路散射波束

现在考虑实际接收到的标准化信号 $S_R(t)$，它是式(2.16)的实部，如下式所示：

$$S_R(t)=\sum_{k=1}^{L}a_k\cos[\omega_c(t-t_0)+\phi_k] \tag{2.17}$$

瞬时局部平均接收功率 P_R 与 $S_R^2(t)$的时间平均值成正比，从式(2.17)中可以看出

$$P_R=\frac{c}{2}\sum_{k=1}^{L}a_k^2 \tag{2.18}$$

其中 c 是一个比例常数，是由式(2.4)给出的精确功率值，它包含了传播距离、阴影衰落和多径衰落 3 种效应，现在主要关注最后一种效应。利用三角代数将式(2.17)展开

$$\begin{aligned}S_R(t)&=\sum_{k=1}^{L}a_k\cos\phi_k\cos\omega_c(t-t_0)-\sum_{k=1}^{L}a_k\sin\phi_k\sin\omega_c(t-t_0)\\&=x\cos\omega_c(t-t_0)-y\sin\omega_c(t-t_0)\end{aligned} \tag{2.17a}$$

其中，有

$$x\equiv\sum_{k=1}^{L}a_k\cos\phi_k\qquad 和\qquad y\equiv\sum_{k=1}^{L}a_k\sin\phi_k \tag{2.19}$$

需要注意到，由中心极限定理，当 L 值很大时，随机变量 x 和 y 服从高斯分布。特别地，因为 a_k 和ϕ_k 是独立随机变量，这些ϕ_k 都假定服从均值为零的均匀分布，则可得到这些随机变量的阶矩如下，具体推导由读者完成。

$$E(x)=E(y)=0,\qquad 且\ E(x^2)=E(y^2)=\frac{1}{2}\sum_{k=1}^{L}E\left(a_k^2\right)\equiv\sigma_R^2$$

高斯随机变量 x 和 y 均值都为零，且有相同的方差 σ_R^2，由 $\cos\phi_k$ 和 $\sin\phi_k$ 的正交性易知 $E(xy)=0$。这里只考虑了短距离(波长级)多路分量的随机变量，而较长距离的变化本质上将其归因于阴影衰落。

现在回到基本的瞬时功率等式式(2.4)，可以将其写成 $P_R=\alpha^2p$，p 在式(2.7)中曾给过以分贝为单位的定义，是由于阴影衰落而产生的局部平均功率项。只考虑多径、短距、随机变化(比如不考虑较长距离的阴影衰落变化)，由式(2.18)和 σ_R^2 的定义，可得

$$c\sigma_R^2=E(P_R)=E(\alpha^2)p=p \tag{2.20}$$

这里用 $E()$ 来表示只对多径项求平均，得到局域平均功率 p。对此功率进一步求平均即可得到平均接收功率 $\bar{P}_R$，这也就验证了之前所说的 $E(\alpha^2)=1$。由阴影衰落功率密度函数表达式(2.9)也可得以 dB 分贝为单位的平均功率 p_{dB} 的平均值。

我们继续分析，由于高斯不相关变量是相互独立的，所以式(2.19)中定义的 x 和 y 也是相互独立的，再利用简单的三角代数，将式(2.17a)重写如下：

$$S_R(t)=a\cos[\omega_c(t-t_0)+\theta] \tag{2.17b}$$

其中，$a^2=x^2+y^2$，$\theta=\text{arctg}\ y/x$，于是 $E(a^2)=E(x^2)+E(y^2)=2\sigma_R^2$。注意到 a 是信号 $S_R(t)$ 的(随机)包络或者幅度，假设 x 和 y 是零均值高斯变量，则信号包络 a 服从瑞利分布(Schwartz, 1990)①如下式：

$$f_a(a)=\frac{a}{\sigma_R{}^2}e^{-a^2/2\sigma_R^2} \tag{2.21}$$

注意到该式是式(2.15)的明确表达，将其中 σ_r^2 替换成了 σ_R^2。特别地，通过比较式(2.17)、式(2.17a)、式(2.17b)和式(2.18)，对 $S_R^2(t)$ 求时间的平均，可得

$$P_R=ca^2/2=\alpha^2p \tag{2.22}$$

c 仍然是一个比例常数，由于 $\alpha=\sqrt{\frac{c}{2p}}a$ 且 a 服从瑞利分布，易知 α 也一定服从瑞利分布，如式(2.15)所示。注意到对式(2.22)的 α 求平均可得式(2.20)，而式(2.20)显示出幅度随机变量 α 和局部平均接收功率 p 之间的关系。

式(2.22)和式(2.15)使我们能够得到瞬时功率 P_R 的概率分布。由简单概率理论，可得

$$f_{P_R}(P_R)=f_\alpha\left(\sqrt{\frac{P_R}{p}}\right)\left|\frac{d\alpha}{dP_R}\right|=\frac{1}{p}e^{-P_R/p} \tag{2.23}$$

瞬时接收功率服从指数分布，其均值为局部平均功率 p。指数分布的略图如图 2.11 所示。

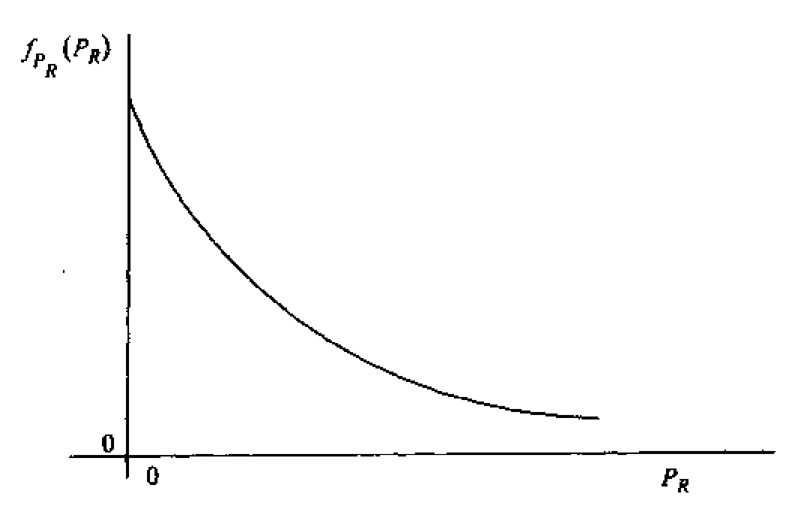


图 2.11 指数分布

① Schwartz, M. 1990. *Information Transmission, Modulation, and Noise*, 4th edn, New York, McGraw-Hill.

由上面可知，当散射波束数 L 较大时，由中心极限定理可知随机变量 x 和 y 为高斯统计量，但是 L 需要多大呢？许多年前的研究显示，只要 6 个有独立随机相位的正弦波相加，所得结果的幅度就近似服从于瑞利统计(Schwartz，Bennett 和 Stein，1966)①。正是这个原因，使这里所讨论的宏小区无线系统中多径衰落模型相当精确。在微小区系统中，我们将要讨论的莱斯分布是一个相对更正确的模型。此系统中，接收机和发射机之间距离更短，多路波束中更有可能有某路波束直接到达接收机，从而在接收中占优势。接下来我们将看到用于获得莱斯统计量的模型。

再看图 2.10，由于在两个天线之间的传播途中遇到障碍发生散射，于是在发射机和接收机之间有多路波束传播。这里假定两个天线之间的距离不是很大，比如在微小区系统中两根天线之间为视距，所以电磁波的一部分能量被直接接收，即称为直接波束的信号。令代表这路直接波束的标准化信号的幅度为 A(但是要注意，这里暗中假设了阴影衰落的情况：直接波束的幅度 A 包含了阴影衰落的影响)。在瑞利情况下，由于散射过程，波束由随机分布幅度和随机变化相位所描述，现在实际接收信号 $S_R(t)$ 可以被写为

$$
\begin{aligned}
S_R(t) &= A\cos\omega_c(t-t_0)+\sum_{k=1}^{L}a_k\cos[\omega_c(t-t_0-\tau_k)+\theta_k] \\
&= (A+x)\cos\omega_c(t-t_0)-y\sin\omega_c(t-t_0) \\
&= a\cos[\omega_c(t-t_0)+\theta]
\end{aligned}
\tag{2.24}
$$

正如式(2.16)和式(2.17)中所做过的一样，但是加上了额外的载波项以体现视距直接波束。随机变量 x 和 y 在式(2.19)中已经被精确地定义了，它们是独立的高斯变量，方差为 σ_R^2。由于相位是相对的量，为简单起见，所有接收直接波束的相位项均置为零。

注意到现在的随机幅度 a 和瑞利情况中的定义有所不同，这里 $a^2=(A+x)^2+y^2$。由 x 和 y 两个独立高斯随机变量进行与瑞利情况类似的分析，可得随机幅度 a 的概率分布。如下为莱斯分布(Schwartz，1990)：

$$
f_a(a)=\frac{a}{\sigma_R^2}\mathrm{e}^{-\left(\frac{a^2+A^2}{2\sigma_R^2}\right)}I_0\left(\frac{aA}{\sigma_R^2}\right) \tag{2.25}
$$

$I_0(z)$是第 1 类零阶贝塞尔函数，定义如下：

$$
I_0(z)\equiv\frac{1}{2\pi}\int_0^{2\pi}\mathrm{e}^{z\cos\theta}\mathrm{d}\theta \tag{2.26}
$$

莱斯分布的略图如图 2.12 所示。与瑞利分布比较可知，瑞利分布是莱斯分布在 $A=0$ 的情况下的特例。易看出，在 $A^2/2\sigma_R^2\gg 1$ 的情况下，莱斯分布逼近于均值为 A 的高斯分布，这也可由式(2.24)预料到。由于对于足够大的 A，含有 y 的积分项可以被忽略不计，于是可得到 $S_R\approx((A+x)\cos\omega_c(t-t_0))$，因此幅度 a 即 $A+x$，是均值为 A 的高斯随机变量。

这些叙述由修正贝塞尔函数 $I_0(z)$的性质不难得到量化。从式(2.26)可以清楚地看到，当

① Schwartz, M., W.R. Bennett, and S.Stein. 1966. *Communication Systems and Techniques*, New York, McGraw-Hill; reprinted, IEEE Press, 1996.

$z \ll 1$ 时，$I_0(z) \to 1$。这对应着式(2.25)中 $A^2/2\sigma_R^2 \ll 1$，导致了瑞利分布。现在考虑其他的极限，当 $z \gg 1$ 时，可能有 $I_0(z) \to e^z/\sqrt{2\pi z}$。利用这一贝塞尔函数的形式，带入到式(2.25)中，并注意到 $A \gg \sigma_r$，此时易知 $f_a(a)$ 在 A 附近到达最高点，并且其幅度值逼近于高斯分布，从而验证了上面的结论。在下一节讨论接收机移动性的影响时，由于小尺度多径变化，我们将回到瑞利统计模型。

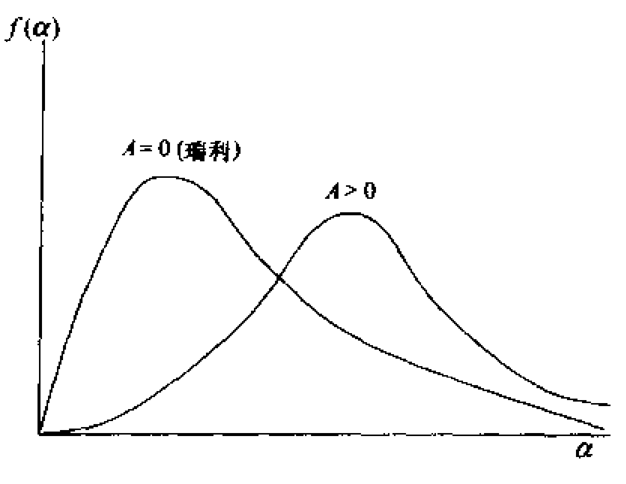


图 2.12　莱斯分布

可由式(2.25)看出，瞬时功率 P_R 的概率分布关于局部平均功率 p 变化，利用 P_R 和信号包络或幅度 a 的关系式(2.22)，可得到在莱斯模型情况下瞬时功率的概率分布如下，推导由读者完成。

$$f_{P_R}(P_R) = \frac{(1+K)e^{-K}}{p} e^{-\frac{1+K}{p}P_R} I_0\left(\sqrt{\frac{4K(1+K)}{p}P_R}\right) \tag{2.27}$$

参数 $K \equiv A^2/2\sigma_R^2$ 称为莱斯 K 因子。在由式(2.25)和式(2.22)推导上式的过程中用到了 $a^2 = (A+x)^2 + y^2$ 和 $E(a^2) = A^2 + 2\sigma_R^2 = 2\sigma_R^2(1+K)$。莱斯 K 因子与所有散射波束平均接收视距或直接波束的信号能量和平均接收信号能量的比值紧密相关。在微小区环境下的测量显示出 K 从 6 dB 到 30 dB 范围内的衰落。

2.3　随机信道特性

在对信号衰落进行更进一步的讨论之前，先更深入地介绍一下发射信号要经过的无线媒介的特性。这一特性将对理解数字信号通过信道传输很有帮助，另外，在本章的后面也可看到这一特性对于减轻一些信号传输过程中的衰落效应也很有帮助。注意到在写出如式(2.16)所示的表达式时，我们曾指出载波信号在通过常称为无线信道的无线媒介进行传输时，由于在传输路径中发生散射，而得到随机多径信号，如式(2.16)所示的散射引入了随机幅度、相位和时延项。在更一般的情况下，发射的是携带信息的调制后的载波信号，这一更普遍的发射信号的复相量形式可表示如下：

$$\tilde{S}_T(t) = s(t)e^{j\omega_c t} \tag{2.28}$$

其中，$s(t)$ 表示调制(携带信息)信号的复相量，调制信号也常称为基带信号。信号的复相量表

示 $s(t)$也常称为信号的复包络。假定 $s(t)$的基带带宽为 $B/2$,则发射载波信号 $\tilde{S}(t)$的带宽为 B,是基带带宽的两倍。式(2.16)接收信号的一般表示可写为

$$\begin{aligned}\tilde{S}_R(t) &= \sum_{k=1}^{L} a_k s(t-t_k) e^{j[\omega_c(t-t_o-\tau_k)+\theta_k]} \\ &= a(t) e^{j\omega_c(t-t_0)}\end{aligned} \tag{2.29}$$

这里 $a(t)$代表接收基带信号的复幅度,表达式如下:

$$a(t) = \sum_{k=1}^{L} a_k s(t-t_k) e^{j(\theta_k-\omega_c\tau_k)} \tag{2.30}$$

式(2.29)是对早些提到的式(2.17b)的推广,式(2.17b)是未调制信号的接收信号,且为实信号形式表示。在隐含多径散射的式(2.16)和式(2.29)中,由散射引入的幅度和相位项并不随时间变化,于是无线信道可以被建成一个线性时不变模型。更一般地,如果幅度和相位随时间变化,则信道为线性时变信道。举例来说,考虑接收机移动的话,就将把时变引入到信道模型中来。下面对时不变情况和时变情况都将进一步讨论。

线性时不变信道模型可由式(2.29)和式(2.30)的比较直接引出,特别注意到接收和发射基带信号,分别为 $a(t)$和 $s(t)$,对于一般线性时不变信道模型可易知输出信号 $a(t)$是输入信号 $s(t)$和信道冲激函数 $h(t)$的卷积

$$a(t) = \int_{-\infty}^{+\infty} h(\tau) s(t-\tau) \mathrm{d}\tau \tag{2.31}$$

此情况下的冲激响应 $h(t)$是基带的(随机)信道响应。对于这里定义的散射信道,比较式(2.31)和式(2.30),可得 $h(\tau)$的离散信号表示

$$h(\tau) = \sum_{k=1}^{L} a_k e^{j(\theta_k-\omega_c\tau_k)} \delta(t-\tau_k) \tag{2.32}$$

式(2.32)中的函数 $\delta(t-\tau_k)$是一般的单位冲激,或者是除了在 $t=\tau_k$ 处其值为 ∞,其他时间均为零的 δ 函数,且满足$\int\delta(t)\mathrm{d}t=1$。

现在考虑由 L 路散射体引入的幅度、时延和相位关于时间变化的情况。如上所述,这一时变可能是由于下节将讨论的终端移动引起的,也可能是由散射体本身移动引入的,后者的例子如散射体是大型卡车、公共汽车或者火车。在幅度、时延和相位随时间变化的情况下,在式(2.29)和式(2.30)中将有 $a_k(t)$、$\tau_k(t)$和 $\theta_k(t)$。无线信道时变形式的卷积如下:

$$a(t) = \int h(t,\tau) s(t-\tau) \mathrm{d}\tau \tag{2.33}$$

冲激响应 $h(t,\tau)$是时变的,并且是一个随机过程,在任一时间其值均为随机变量(Papoulis, 1991)①。结合目前考虑的情况,线性时变信道的冲激响应为

$$h(t,\tau) = \sum_{k=1}^{L} a_k(t) e^{j[\theta_k(t)-\omega_c\tau_k(t)]} \delta(t-\tau_k(t)) \tag{2.34}$$

---

① Papoulis, A. 1991. *Probability, Random Variables, and Stochastic Processes*, 3rd edn, New York, McGraw-Hill.

冲激响应 $h(t,\tau)$是两个时间变量 t 和 τ 的函数,定义为在时刻 t 之前 τ 个单位时间,也就是 $t-\tau$ 时刻的冲激在时刻 t 得到的响应。在规划 GSM 蜂窝系统时,对其进行了研究。我们将在后面章节详细讨论各种环境下使用不同的信道冲激响应模型来估计系统性能,包括乡村地区,丘陵环境和城市地区(Steele,1992)①。为此目的采用了 6 部分和 12 部分模型,在这些模型中考虑了接收机移动的影响。在下节我们将看到,这一运动将由接收频率的多普勒频移表现出来。

随机过程 $h(t,\tau)$是两个变量的函数,这是由于它表现了时变线性信道的特性。在 20 世纪 60 年代初,Bello 对随机无线信道中这样的随机过程进行了详尽的研究(Bello,1963)②。为简化研究,它假定这些信道是广义平稳信道,它们的一阶和二阶矩都是和时间无关的,这一假设大大简化了这些过程的研究(所有阶矩都和时间无关的随机过程称为狭义平稳过程(Papoulis,1991),广义平稳过程的条件要弱一些)。为了理解无线信道中广义平稳信道的假设,首先需要讨论一下只有一个随机变量的广义平稳随机过程的意义。假设一个随机过程 $x(t)$随时间变化,在任一时间下其取值为一个概率分布函数为 $F_x(x)$的随机变量。对随机过程 $x(t)$的一个测度就是它的自相关函数 $R_x(t_1,t_2)$,定义为在两个不同时刻 t_1 和 t_2 的 $x(t)$乘积的总体均值,即

$$R_x(t_1,t_2)=E[x(t_1)x(t_2)] \tag{2.35}$$

其中,E 或者说求均值操作是对括弧里的乘积求均值,这一相关函数粗略地测量了两个不同时刻 t_1 和 t_2 的 $x(t)$取值之间的相关性。特别地,如果两个量不相关,则可得到自相关函数 $R_x(t_1,t_2)=E[x(t_1)]E[x(t_2)]$,即两个均值的乘积。如果这两个值都是零,那么(不相关的)自相关函数是零(经常拿随机变量减去均值来定义相关函数,这样不相关函数的自相关就为零了)。现在考虑一个广义平稳过程,它的自相关函数定义为不随时间变化,只与所计算的时间的差值相关,而与具体的时间无关,其均值 $E(x)$对于时间是个常数,于是这一广义平稳过程的自相关函数为

$$R_x(\tau)=E[x(t+\tau)x(t)]=E[x(t)x(t-\tau)]=R_x(-\tau) \tag{2.36}$$

这里 $\tau=t_1-t_2$,是两个时间的差值。如果随着时间差的减少,相关函数逼近均值的乘积,则有均值自乘,随机变量 $x(t+\tau)$和 $x(t)$称为不相关。对于零均值的取值,在时间上足够大的间距内自相关函数都为零。注意到两个不相关的随机变量不一定是相互独立的,独立性的要求条件更强。不过高斯随机变量如果不相关的话,就是相互独立的。相反地,所有相互独立的随机变量都是不相关的。功率谱密度是随频率变化的功率随机过程的一种测度,而广义平稳随机过程的功率谱密度是其自相关函数的傅里叶变换(Papoulis,1991)。

自相关函数的定义可扩展到时变随机过程中,例如,式(2.34)所示的时变无线信道的冲激响应 $h(t,\tau)$的复相量表达式。特别地,假定一个广义平稳随机信道,可得到

$$R_h(\tau,\tau_1,\tau_2)=E[h(t+\tau,\tau_1)h^*(t,\tau_2)] \tag{2.37}$$

操作符 * 表示对 h 求复数共轭,在这种情况下时变信道的自相关函数是一个有 3 个参数的函数:它代表了两个冲激响应乘积的均值,一个是测量时间 $t+\tau$,另一个是 t,提前了 τ 个单位时间,而且在 WSS 过程式(2.36)中,自相关只与这两个测量时间的差值 τ 有关,而与具体的实际

---

① Steele, R. ed. 1992. *Mobile Radio Communications*, London, Pentech Press; New York, IEEE Press.

② Bello, P. A. 1963. "Characterization of randomly time-variant linear channels," *IEEE Transactions on Communication Systems*, 12, CS-11 (December), 360 – 393.

计算时间无关。在时间 $t+\tau$ 和 t 的冲激响应分布为在此时间之前 τ_1 和 τ_2 的冲激函数。在计算自相关函数时，这 3 个时间 τ,τ_1 和 τ_2 均需要考虑在内。由式(2.34)所表示的 $h(t,\tau)$，是由 L 个散射体的多径模型得来的，是 WSS 过程的一个特例，并且它由 Bello 定义来作为广义平稳不相关散射过程(WSSUS)的一个例子(Bello，1963)。

在下节中我们将重点介绍终端的移动对式(2.16)和式(2.17)的多径接收信号模型的影响，这将需要讨论移动终端所经历的衰落速率。在 2.5 节中引入了由式(2.28)表示的发射信号的简单双频模型，这使得我们可以对多径环境下的频率选择性衰落现象进行介绍。

2.4　终端移动和衰落速率

截止到目前的讨论，一直都没有考虑终端的移动性，在本节中将考虑这一点，并讨论其对信号衰落产生的影响。特别地，假设一辆车或者一个人带着一个移动终端在 x 轴方向上以速度 v km/h 移动。假设到达移动终端的传输波都被垂直地极化，且到达移动终端的第 k 路波与水平方向成夹角为 β_k，如图 2.13 所示。对于接收波，车辆的移动引入了一个多普勒频移 $f_k = v\cos\beta_k/\lambda$ Hz。如果 $\beta_k=0$，也就是说波束到达方向与车辆移动方向刚好相反，则有 $f_k = v/\lambda$，且接收信号的频率增加为 f_c+v/λ，其中 f_c 为未调制载波频率。如果 $\beta_k=\pi$，也就是说波束到达方向与车辆行驶方向相同，那么接收信号的频率减少为 f_c-v/λ。这些值均为考虑了多普勒速度后得到的结果，但是也注意到这些频移都很小。举例来说，即使对一个 100 km/h(60 mile/h)的车辆和频率为 1 GHz 的系统而言，$\lambda=0.3$ m，$v/\lambda=90$ Hz，然后得到$(v/\lambda)/f_c\approx10^{-7}$！一个人携带一部手机行走，其速度为 5 km/h，这意味着产生了 v/λ 为 4.5 Hz 的频移。尽管这些多普勒频移都很小，但是它们对蜂窝系统的性能和接下来要讨论的衰落速率都起着重要的作用。

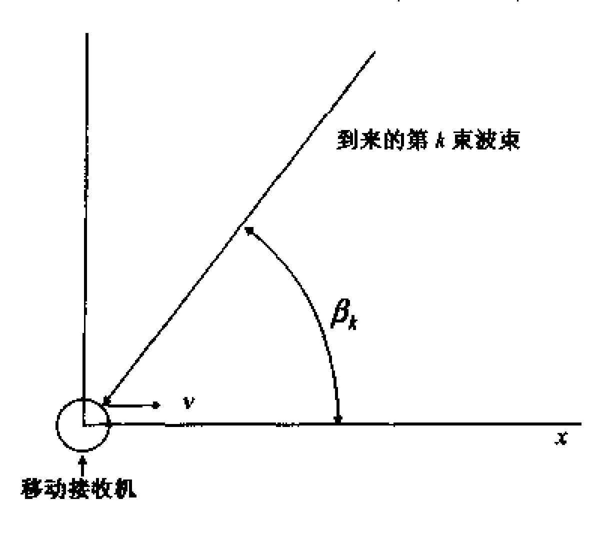


图 2.13　多普勒效应的几何图

由于终端的移动产生的多普勒频移，将会使接收到的多路信号产生一个附加相移。对于第 k 路，其附加相移为 $2\pi f_k t=\omega_k t=(2\pi v\cos\beta_k)t/\lambda$。注意到式(2.17)给出的瑞利多径衰落情形，在终端移动的情况下，接收信号 $S_R(t)$ 可表示为

$$S_R(t)=\sum_{k=1}^{L}a_k\cos[\omega_c(t-t_0)+\phi_k+\omega_k t] \tag{2.38}$$

可以利用对瑞利统计量和多普勒频移的讨论，来对衰落速率、多径效应和传输信号的多普勒频移进行定量的说明。先看衰落，已经指出信号的衰落速率和衰落深度是与这两项相关的：一个是由小尺度多径现象引出的瑞利模型，另一个是移动终端的速率。由于衰落是由连续多径波束的有益和有害干扰叠加而成的，距离上在波长级范围内变化，所有衰落速率的变化一定与 v/λ 有关。举例来说，如果一辆车以速度 $v = 100$ km/h = 28 m/s 行驶，发射信号的频率 f_c 为 1 GHz 或者 $\lambda = 0.3$ m，与前面所举的例子一样，那么衰落速率应该在 $v/\lambda = 90$ fades/s 左右，而衰落的持续时间是 λ/v 级的，或者是 11 ms。对于一个人带着手机以 5 km/h 行走，衰落比车辆要慢一个因子 20，衰落速率应该是以 4.5 fades/s 为步长，且衰落持续时间以 0.2 s 为步长。可以利用瑞利统计量使其更精确，特别地，对于垂直极化的电场，在移动终端测得的平均衰落持续时间 τ_f 如下式所示(Jakes,1974)：

$$\tau_f = \frac{e^{\rho^2} - 1}{\rho f_m \sqrt{2\pi}} \tag{2.39}$$

这里 $f_m \equiv v/\lambda$ 是最大的多普勒频移，且 $\rho \equiv a/\sqrt{E(a^2)}$，$a$ 是瑞利分布中接收信号的幅度。所以，平均衰落持续时间与测得的信号幅度相关，参数 ρ 是幅度 a 对其均方根 $\sqrt{E(a^2)}$ 的标准化。Jakes(1974)和 Yacoub(1993)都各自推出过式(2.39)。作为一个例子，假定幅度值就是它的均方根值，那么就有 $\rho = 1$，且由式(2.39)可得平均衰落持续时间 $\tau_f \approx 0.7/f_m = 0.7\lambda/v$，平均衰落速率为 $1/\tau = 1.4v/\lambda$。这些值都和直观上得到的值很接近。对于上面提到的以速度 100 km/h 行驶的车辆和工作频率为 1 GHz 的例子，如果其幅度与均方根比值 $\rho = 1$，那么其平均衰落持续时间大约为 $\tau_f = 8$ ms；如果 $\rho = 0.3$，也就是说接收到的信号幅度 a 是其均方根值的 0.3，那么平均衰落持续时间就降为 $0.1/f_m$ 或者 1 ms。对于同一个例子，如果移动速度为 5 km/h，则这两个数将分别增加为 0.12 s 和 20 ms。

我们是如何得到式(2.39)的呢？这个分析相当长，所以只概括一下使用的方法，具体的细节可参阅 Yacoub(1993)的著作。假定一个特殊的移动接收机接收到的信号幅度 $a = R$，该振幅的平均穿越率 R_c 如下式所示：

$$R_c \equiv E[\dot{a}/a = R] = \sqrt{2\pi} f_m \frac{R}{\sqrt{2}\sigma_R} e^{-\frac{R^2}{2\sigma_R^2}} \tag{2.40}$$

f_m 是最大多普勒频移 v/λ；σ_R^2 是前面提过的瑞利分布参数[$E(a^2) = 2\sigma_R^2$]。图 2.14给出了一个典型的时间函数的衰落形式，从图中可以看到幅度 R 及其所决定的穿越率(斜率)R_c。对于 $R \ll \sigma_R$，幅度穿越率很小；对于 $R > \sigma_R$，幅度穿越率也很小，所以易看出 R_c 在幅度 $R = \sigma_R$ 时到达最大值 $1.08f_m$。

平均衰落持续时间 τ_f 的表达式可由式(2.40)如下获得。考虑一个长时间间隔 T，假定信号幅度衰落或下降在电平 R 以下，在此时间间隔内发生了 n 次，第 i 次衰落持续时间为 τ_i s，在图 2.14 中举了衰落持续时间 τ_1 s 的一个例子。将这些衰落时间相加，并对 T 求标准化，则 $\sum_i \tau_i / T$ 表示幅度值在电平 R 以下的时间比率。对于大的 n，这个量是电平衰落到 R 以下概率的一个很好的近似，这也正是瑞利分布信号幅度 a 比 R 小的概率，

$$P(a \leqslant R) = \int_0^R f_a(a) da = 1 - e^{-\frac{R^2}{2\sigma_R^2}} \tag{2.41}$$

现在，注意到$\sum_i \tau_i / T$也可以写做$(\sum_i \tau_i / n) n / T$，其中$n/T$正是穿越电平$R$水平线和在大$T$下到达平均穿越率$R_c$的次数，同时$\sum_i \tau_i / n$逼近平均衰落持续时间$\tau_f$。于是，可由式(2.41)得到

$$\tau_f R_c = 1 - e^{-\frac{R^2}{2\sigma_R^2}} \tag{2.42}$$

此时，可由式(2.40)的R_c和式(2.42)得到平均衰落持续时间τ_f，即可得到想要的式(2.39)。

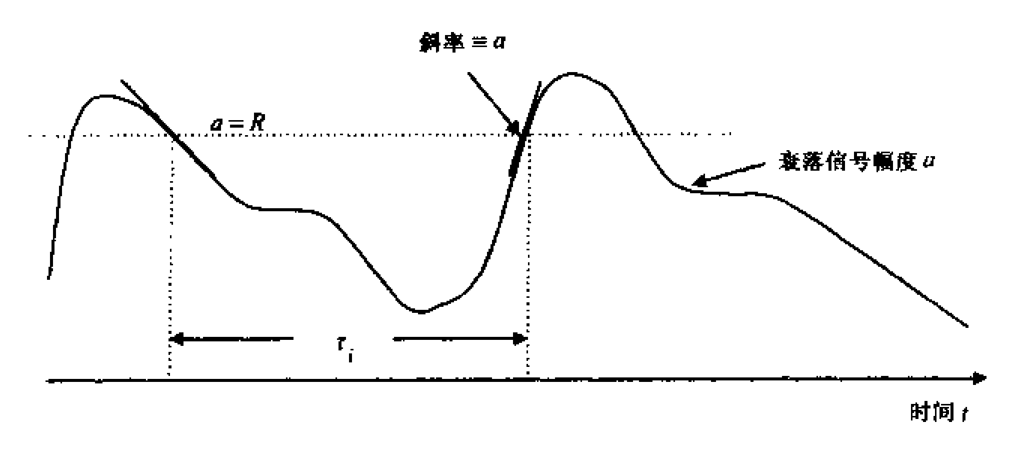


图 2.14　衰落信号和幅度穿越

2.5　多径和频率选择性衰落

上一节主要讨论了信号衰落、衰落时的速度和其倒数衰落的持续时间，它与终端移动速度有关，如多普勒频移参数$f_m = v/\lambda$所示。这一讨论也集中于单载波未调制发射波的简单情况。实际通信中，接收和发射信号都是携带信息的调制载波信号。在 2.3 节信道特性中，修正了 2.2 节关于多径接收的讨论，发射信号$s(t)$有相应的基带带宽$B/2$，这时发射信号的复相量表示由式(2.28)给出。我们也注意到发射带宽是基带带宽的两倍即B Hz。在这一讨论中暗含了幅度调制的假设，如式(2.28)所示。在移动无线通信系统中，所用的其他类型的信号调制将在第 5 章中讲述。在 FM 的情况下，发射带宽是基带带宽较复杂的函数(Schwartz, 1990)。现在来看发射携带信息的信号这一更实际的传播模型。图 2.15(a)给出了一个典型的发射信号的频率谱，是一条幅度随频率变化的曲线，从中可看到在载波f_c附近的发射带宽为B。

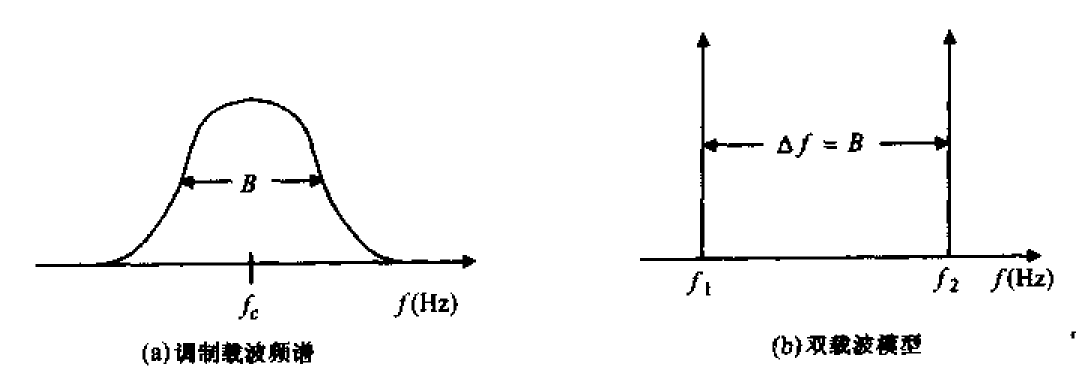


图 2.15　信号带宽

这一节将讨论多径衰落对接收携带信息的信号的影响，这一影响与信号带宽有关。对于大的带宽将会发生频率选择性衰落，于是在传播信道上信号不同的频率成分有了不同的变化，

这就导致了信号的失真。对于本书讨论的数字信号情况,频率选择性衰落引入的失真由其自身的符号间干扰(ISI)表示,它是由相继的数字符号与相邻符号部分交叠而成的。对于窄带宽信号,发生的是无选择性衰落或者是常说的平衰落。在相反的极端情况即非常宽的带宽情况下,同一信号的多径复制信号可能在明显不同的时间到达接收端,如果允许分开到达的信号可以单独处理,那么对原始信号的检波能力将会提高。这一现象实际上可用于提高 CDMA 信号的性能上,这点将在本书的后面介绍,在本章的最后也将简要介绍一下这一现象。

为研究传输路径对信号接收的影响,用可能的、最简单的调制载波模型式(2.28),同时发射(和接收)两路正弦波。这两路有着相同发射功率,但在频率上有少许的差别,图 2.15(a)给出了实际携带信息的信号的频率差异,如图所示它们之间的差值即为带宽。利用这一简单的模型来确定多径衰落信道对两路信号的作用,从而近似知道实际信号到达时是否由于信道而发生失真,多径成分是否能被接收机处理,终端移动对接收信号的影响是否能一样被处理。尤其是最后的这一影响,在上一节中使对衰落速率的讨论得到了进一步的展开。特别地,考虑两个信号,信号 1 和信号 2,信号 1 的载波频率为 f_1,信号 2 以稍高的频率 f_2 被发射,两频率差值 $\Delta f = f_2 - f_1$。在实际信号中,这一差值与带宽 B 相等。这一情况如图 2.15(b)所示[当提到频率间稍微不同时,意味着其波长很接近于同一值,即 $\Delta\lambda = \lambda_2 - \lambda_1 \ll \lambda_1$(或 λ_2),这也意味着 $\Delta f \ll f_1$(或 f_2)]。

现在假设这两路发射信号在传播路径上各自发生散射,到达移动接收机时是大量多径波束的和。为了使模型一般化,假设接收的两路散射波是由不同数目的多径波束组成的,同样这两组波束有不同的传播时延。在上节式(2.38)中,考虑了由于终端移动引入的多普勒频移,将其进行扩展,可得到

$$S_1(t) = \sum_{k=1}^{L} a_k \cos[\omega_1(t - t_1 - \tau_k) + \omega_k t + \theta_k] \tag{2.43a}$$

和

$$S_2(t) = \sum_{l=1}^{M} a_l \cos[\omega_2(t - t_2 - \tau_l) + \omega_l t + \theta_l] \tag{2.43b}$$

如先前所提到的[见式(2.17)和式(2.24)],t_j 表示各个多径波束的大尺度时延,其中 $j=1$ 和 2;τ_k 和 τ_l 表示对每组波束中每一个波附加或增加的随机时延;$\omega_k = 2\pi\cos\beta_k v/\lambda$ 和 $\omega_l = 2\pi\cos\beta_l v/\lambda$;$\theta_k$ 和 θ_l 是附加的随机角度,假设它们服从 0 到 2π 的均匀分布。数字 M 和 L 分别是两路发射信号到达接收端的各自的波束数目。β_k 和 β_l 表示每一个波到达方向与水平面的夹角,如图 2.13 所示。假设大尺度时延 t_1 和 t_2 相等,我们将得到调制载波带宽为 Δf 的模型。每一个频率项加上随机时延,将决定无线信道的频率选择性。保留大尺度时延使我们可以看出移动速度和这些时延之间的关系。

为继续讨论,必须为式(2.43a)和式(2.43b)附加的时延定义一个模型。一个常用的模型是假设这些时延服从均值为 τ_{av} 的指数分布,于是较大的附加时延的概率以指数规律降低,附加时延的概率密度函数 $f_\tau(\tau)$ 为

$$f_\tau(\tau) = \frac{1}{\tau_{av}} e^{-\frac{\tau}{\tau_{av}}}, \quad \tau \geqslant 0 \tag{2.44}$$

这一函数的略图如图 2.16(a)所示，易知它的二阶矩 $E(\tau^2)$ 为 $2\tau_{av}^2$，所以方差 $\sigma_\tau^2 \equiv E(\tau - \tau_{av})^2 = \tau_{av}^2$。这一附加时延分布的标准差 $\sigma_\tau = \tau_{av}$ 被定义为时延扩展。如图 2.16(b)所示，各个接收信号总的时延为随机变量 $t_i + \tau$。宏小区系统时延扩展的典型值，在乡村地区为 0.2 μs，在郊区为 0.5 μs，在城市地区为 3 ~ 8 μs，这是由于城市地区有更多的障碍物，所以会遇到更多的多径(Yacoub，1993[①]；Steele，1992；Rappaport，2002[②])。城市微小区系统的时延扩展小于 2 μs，而(市内)微微小区系统的时延扩展为 50 ~ 300 ns。

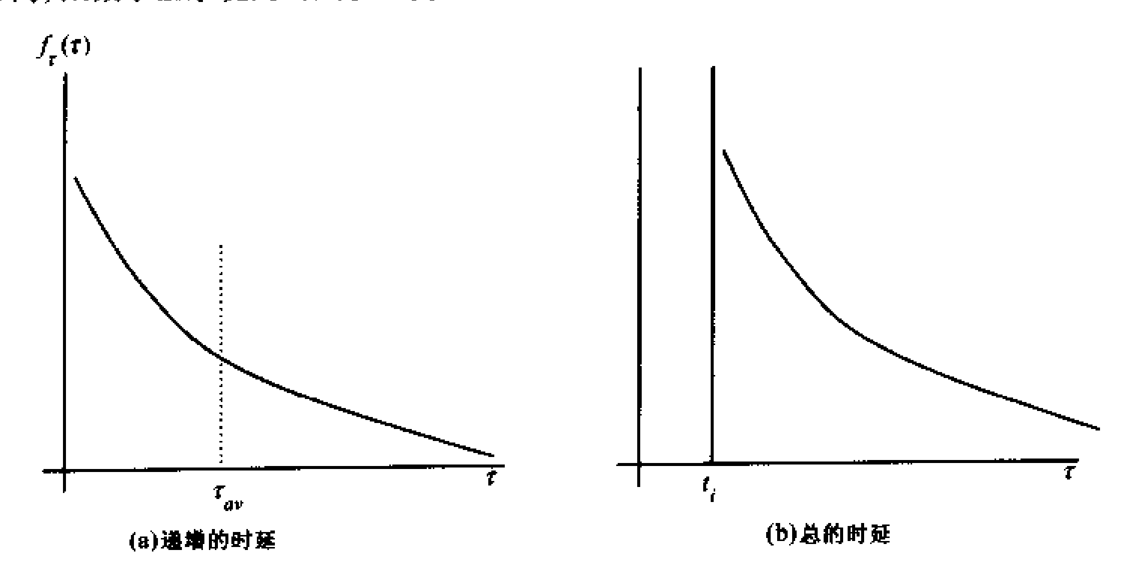


图 2.16　指数时延分布

式(2.43a)、式(2.43b)和式(2.44)可用于确定接收机移动性和携带信息的调制信号衰落的联合影响，为此用式(2.17b)中包络/相位的形式重写式(2.43a)和式(2.43b)。不过，我们将其写成 $S_i(t) = a_i\cos(\omega_i t + \theta_i)$，$i = 1,2$ 的形式，其中传播时延为 t_1 和 t_2，载波频率为 ω_1 和 ω_2，将其分别相乘并加入各自的相位项和幅度项。随机变量 a_i 和 θ_i 是附加时延分布和多普勒频移的函数。我们将这两路接收信号 $S_1(t)$ 和 $S_2(t)$ 相比较并表现出衰落信道对单个调制信号影响的方法，就是它们的包络相关函数 $E(a_1a_2)$。对式(2.43a)和式(2.43b)中所有的随机变量 $a_k s$，$a_l s$，$\beta_k s$，$\beta_l s$，附加时延分布和随机相位 θ_k 与 θ_l 求平均，可得到该包络的相关函数。对我们的目的来说，标准化的包络相关函数更有用，该函数定义为 $[E(a_1a_2) - E(a_1)E(a_2)]/\sigma_1\sigma_2$，其中

$$\sigma_i^2 \equiv E[a_i - E(a_i)]^2, \quad i = 1,2$$

是包络 a_i 的方差，其均值为 $E(a_i)$。从现在开始记这一标准化的包络相关函数为 ρ_a，于是有

$$\rho_a \equiv [E(a_1a_2) - E(a_1)E(a_2)]/\sigma_1\sigma_2 \tag{2.45}$$

这两个随机变量的相关，类似于前面讨论的自相关函数，它给出了随机过程中两个相隔 τ 时间的实例之间的相关性(也可以将自相关函数标准化)。在自相关情形下，如果 $E(a_1a_2) = E(a_1)E(a_2)$，那么这两个随机变量认为是不相关的，于是有 $\rho_a = 0$。当 $\rho_a = 1$ 即 ρ_a 为最大值时，这两个随机变量是完全相关的。在这里所讨论的特殊情形下，ρ_a 接近于 1 表明衰落信道

---

① Yacoub, M.D. 1993. *Foundations of Mobile Radio Engineering*, Boca Ration, FL, CRC Press.

② Rappaport, T.S. 2002. *Wireless Communications, Principles and Practice* 2nd edn, Upper Saddle River, NJ, Prentice-Hall.

是不失真的,也就是说调制信号所有的谱分量以同样的方式发生衰落。当 ρ_a 接近于 0 时,有两种情况:信道使携带信息的信号失真,或者如果带宽(在模型中的频率间隔)足够大,可能会出现分离的波束或回波,使它们可以被分别处理。如果单个的回波被适当地重新组合的话,将使性能得到切实的提高。这样的技术包括 RAKE 接收机,我们将在本章的最后简要地讨论。在本章的后面将看到 CDMA 无线系统包含重组接收到的分离波束这一过程,从而提高系统性能。

关于 ρ_a 的计算可在两路信号频率分离的情况下进行,假设瑞利衰落及附加时延为指数分布。在冗长的计算(Yacoub,1993:132 - 138)之后,可得到 ρ_a 相对简单的近似表达式如下:

$$\rho_a(\Delta\tau,\Delta f)\approx\frac{J_0^2(\omega_m\Delta\tau)}{1+(\Delta\omega\tau_{av})^2} \tag{2.46}$$

$J_0(x)$是第 1 类零阶贝塞尔函数;$\Delta\tau=t_2-t_1$ 是两路到达信号的时延(假定信号 2 晚到达);$\Delta\omega=\omega_2-\omega_1=2\pi(f_2-f_1)=2\pi\Delta f=2\pi B$,其中 B 是发射信号带宽,在此模型中也为两路载波的频率差值;τ_{av} 是先前定义的时延扩展;$\omega_m=2\pi f_m=2\pi v/\lambda$ 是最大的多普勒频移,单位为rad/s。图 2.17 给出了标准化相关函数 ρ_a 的略图,图中 ρ_a 为在不同 $\omega_m\Delta\tau$ 值下关于 $\Delta\omega\tau_{av}$ 变化的函数。可以看到,随着 $\Delta\omega\tau_{av}$ 或 $\omega_m\Delta\tau$ 的增加,这两路信号的相关趋于零,即为失真信道,现在解释一下这些结果。

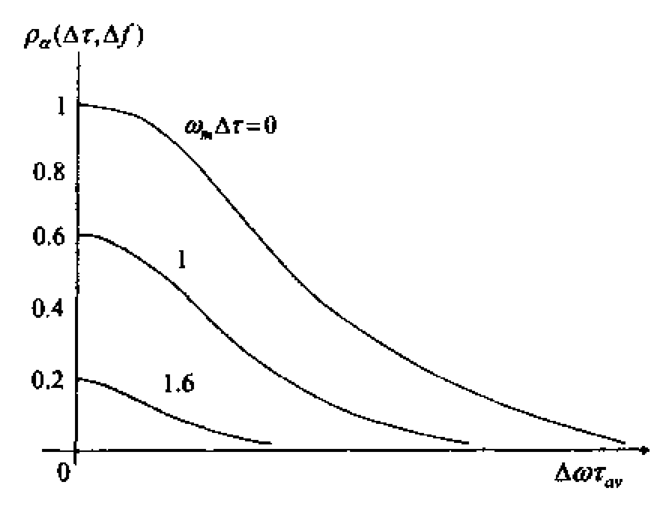


图 2.17　标准化相关函数

首先看一下 $\Delta\tau=t_2-t_1=0$ 的情况。此时,在大尺度意义下两路信号同时到达,知道这两路信号代表的是同一信号的不同频率分量。在此情况下,模拟带宽为 $B=\Delta f$ 的携带信息的信号在多径传播时的效应,就可以确定频率选择性衰落的条件。由于 $J_0(0)=1$,所以有

$$\rho_a\equiv\rho_a(0,\Delta f)=1/[1+(\Delta\omega\tau_{av})^2] \tag{2.47}$$

令 ρ_a 小于或者等于 0.5,则两频率项不相关,从而引出频率选择性衰落。对于 ρ_a 的这个值,$\Delta\omega\tau_{av}\geqslant1$,但是由 $\Delta\omega=2\pi(f_2-f_1)=2\pi\Delta f\equiv2\pi B$,两路信号的频差与携带信息的信号的带宽相等,都为 B Hz。由于 $\rho_a(0,\Delta f)=0.5$,故这一特殊值 $\Delta f=B=1/2\pi\tau_{av}$被定义为相干带宽。这一值被作为频率分离的测度(一个真实携带信息信号的等效带宽 B),当超过这个值时将发生频率选择性衰落。注意到,也可将 $\tau_{av}>1/2\pi B$ 作为频率选择性衰落发生的条件,这将引发前面提到的符号间的干扰。这里最常见的数字符号是二进制数字或比特,在第 5 章中将看到多电

平数字信号的传输,每个符号对应多个比特。下面来看符号间隔 T,在几个单位时间内,当连续的符号在此时间叠加时,发生符号间的干扰。在第 5 章中,我们将看到符号间隔 T 和数字信号发射带宽 B 互为倒数,$BT \approx 1$(这一关系由傅里叶分析得到,所有具备数字通信或通信理论知识背景的人都应该很熟悉)。于是,可得到引发符号间干扰的频率选择性衰落的两个等价条件

$$T < 2\pi\tau_{av} \tag{2.48a}$$

或

$$\tau_{av} > 1/2\pi B \tag{2.48b}$$

当数字符号间隔 T 小于 5 或 6 倍的时延扩展 τ_{av} 时,将有频率选择性衰落和随之发生的符号间干扰,这一结果与直觉吻合。

考虑下列例子:

例 1

数字数据的速率为 10 Kbps,以二进制的形式传输,于是二进制间隔 T 为 0.1 ms,在时延扩展大于或等于 0.02 ms(20 μs)的环境下将发生符号间的干扰。由上述可知,这么大的时延扩展在大多数衰落环境下不会发生,所以大多数衰落媒介对于这种速率传输的数据都只会平衰落或者非频率选择性衰落。

例 2

将数据传输速率提高到 200 Kbps,符号间隔相应地减少为 0.005 ms 即 5 μs,当时延扩展大于或者等于 1 μs 时,将发生频率选择性衰落。由上述可知,这样的时延扩展在城市环境下和一些郊区地区会遇到。在后面的章节中,将看到 GSM 蜂窝系统以 271 Kbps 的速率和 200 kHz 的发射带宽传输信号。

例 3

现在考虑传输速率为 54 Mbps 的情况,在第 12 章中将看到,一些高比特速率无线局域网以这个速率传输信号。频率选择性衰落与之导致的符号间干扰,如果时延扩展大于或者等于 3 ns 时,这样的衰落就会引起问题。这一潜在的问题将在下面两个情形中遇到:在第 5 章的讨论中,当使用 QAM 调制技术时有意地将符号间隔增大到 6 倍;垂直频分复用(OFDM)也是在第 5 章中讨论的,它被用来提供很窄的传输频带,从而在每一路频带上发生平衰落。

这里的讨论主要集中在发生符号间干扰的频率选择性衰落上,现在考虑一个极限情况:信号的带宽 B 足够大,使得时延扩展远远大于 $1/2\pi B$。对应于不同多径波束的不同信号回波将会出现,从而可以对它们分别进行处理。如早些所讲,这一现象可以用来提高在高带宽系统中信号的检波能力。举个例子,假设一个系统的传输带宽 B 为 1.25 MHz,这接近于 IS-95 的带宽,是本书后面要讨论的第二代 CDMA 蜂窝无线标准,与第 10 章要讲的第三代系统 cdma2000 的带宽也一样。然后,当得到的多径回波间距都远远大于 $1/2\pi(1.25)$ μs ≈ 0.13 μs 时可被分解开,单独的回波可以被分别检测和适当的处理,以改善系统性能。我们将在下一节的最后讨论这种称为 RAKE 接收机的处理过程。举第 2 个例子,若在第 8 章中所讲的数字 AMPS 系统 IS-136 的带宽为30 kHz,则意味着同一传输信号的多径波束或者回波间距需要大于 $1/2\pi(30)$ms $\approx$ 5 μs 才能被分离开,这是不太可能达到的(这对应于传播路径有 1000 m 的差距,回忆前面讨论的时延扩展,其值在乡村地区不到 1 μs,在城市地区为 3 μs)。所以,在像 IS-136 的窄带系统中单独的多径回波不能被分离开。在这样的系统中,可以利用信道均衡来减轻频率选择性衰落的影响,在下一节中也将简要介绍信道均衡。

现在回到式(2.46),考虑贝塞尔函数分子的参数,注意到 $\rho_a(\Delta\tau, 0)$ 可以提供有关多普勒频移的标准化相关函数的独立性和时延大尺度差异的信息。贝塞尔函数 $J_0(\beta)$ 如图 2.18 所

示,注意到它是参数 β 的振荡函数,每一次振荡振幅都在不断减小,对于大的 β 将逼近于零。由式(2.46)可得 $\rho_a(\Delta\tau,0)=J_0^2(\omega_m\Delta\tau)$,这一函数的略图如图 2.19 所示,可看到相关性下降得很快,在乘积 $\omega_m\Delta\tau$ 足够大时逼近零。对于上述类似的涉及到 $\Delta\omega\tau_{av}$ 或者等价 $B\tau_{av}$ 的情况,这里的解释如下:当传输信号的不同分量到达时,如果时间上相隔足够远的话,它们会因终端的移动性经历不同的衰落环境。特别地,衰落信道的相干时间,这个由相干带宽类推而来的概念,被定义为当$\rho_a(\Delta\tau,0)=0.5$时的时间差值 $\Delta\tau$,此值即为 $\Delta\tau=9/8\omega_m=9/16\pi f_m=0.18/f_m$。考虑先前所给的例子,速度为 100 km/h 的车辆在传输频率为 1 GHz 的情况下通信,则有 $f_m=v/\lambda=90$ Hz,相干时间为 $\Delta\tau=0.18/f_m=2$ ms。假定一个数字信号在无线系统中传输,若单独的传输符号间隔大于 $\Delta\tau=2$ ms,则将经历不同的衰落环境,引起信号失真。反之,如果信号符号间隔

$$T<\Delta\tau=1/5.6f_m$$

那么信号就不会失真。在这个例子中,如果 $T<2$ ms 或者以大于 500 符号/s 的速率发射,信号将不会失真。对于二进制传输,其间隔为 T,当发射速率大于 500 bps 时将不会发生失真。

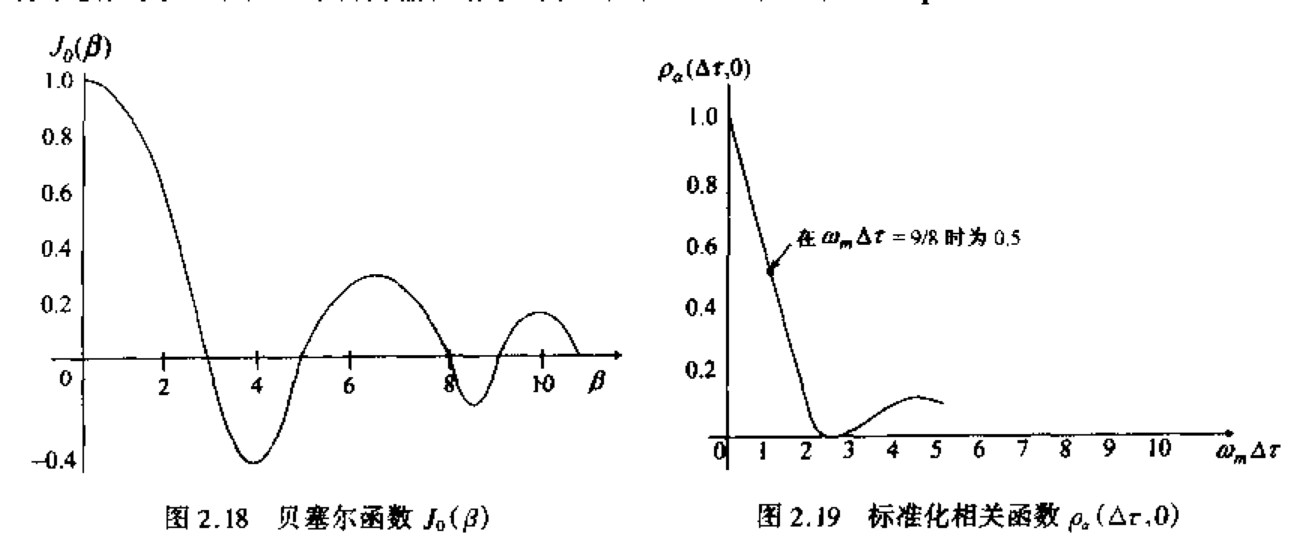


图 2.18 贝塞尔函数 $J_0(\beta)$

图 2.19 标准化相关函数 $\rho_a(\Delta\tau,0)$

我们将扼要重述上面所讲的两类失真来总结关于可能的信号失真的讨论。一个是关于时间色散(符号间干扰)和频率选择性衰落的,当信号带宽超过相干带宽 $1/2\pi\tau_{av}$ 时将会发生这些情况,其中 τ_{av} 是时延扩展。同时也看到,如果带宽足够大的话,多径回波可以被分离开,此时尽管有失真但也可提高信号的检波性能。另外一个是关于接收机移动的影响,如果信号的变化速度与相干时间 $0.18/f_m$ 相比足够快的话,那么信号也不会发生失真,其中 f_m 是多普勒频率 v/λ。多普勒效应引发了频率色散和时间选择性衰落。注意这两个发生在衰落环境下效应的二元性,在 Steele(1992:Chap.2)的著作中有关于这些效应的详细有趣的讨论。第 1 类情况,关于时间色散和频率选择性衰落是由于时延扩展的多径传播造成的,如果信号带宽与时延扩展的倒数相比足够大的话,它的各路波将沿不同路径传播,那么就会发生频率选择性衰落和随之而来的失真和符号间干扰。第 2 类情况是关于频率色散和时间选择性衰落在传输信号过程中如果信道特性改变的话会发生,那么信道特性的改变与接收机的移动直接相关;如果信号本身的变化速度与多普勒扩展 $f_m=v/\lambda$ 的倒数相比足够快的话,那么将不会发生失真[注意到随着信号变化速度的加快,其带宽也将增加(Schwartz,1990)。因此,为消除频率色散和时间选择性衰落,需要一个最小带宽,但是如果信号带宽增加得过大,则会由于时延扩展发生频率选择性衰落]。

这些双重效应发生的条件可由下列等式概括

频率选择性衰落/时间色散

$$B > \text{相干带宽} = 1/2\pi\tau_{av} \tag{2.49}$$

τ_{av}是时延扩展，B 是信号带宽。

时间选择性衰落/频率色散

$$T > \text{相干时间} = 9/16\pi f_m = 0.18/f_m \tag{2.50}$$

f_m 为 v/λ，是最高多普勒频率；T 是信号的符号间隔。

2.6 改善衰落的技术

在上一节中讲述了信号通过无线信道的传输，从中我们可以看到信道环境很苛刻。信号衰落经常存在，需要认真对付；如果数字信号的符号间隔可与时延扩展相比较或者小于它的话，将会发生符号间干扰。经常存在的噪声将更进一步地破坏传输信号。在下一章及以后会详细讨论的信号干扰可能会导致在检测给定传输信号时发生错误。注意到在发送数字信号时，在每一个符号间隔需要发送一些信号中的一个，最简单的例子，二进制传输，每一个二元符号间隔将在两个可能的二进制数字中发送一个。在第 5 章中，讨论 QAM 信号的传输，在每一个符号间隔将发送 $M \geq 2$ 个可能的符号。接收机接收到的符号将用于确定每个符号间隔发射的信号，由于这些影响，一个深衰落的信号更易于被错误地接收。

多种多样的技术被引入来克服衰落、噪声和信号干扰的联合效应；这些效应如果不被处理的话将会导致信号判决错误。数字信号的交织将其在时间上扩展，这一技术一般用于减小快衰落和可能的噪声突发所带来的影响，第 8 章和第 10 章在分别讨论第二代和第三代蜂窝系统时将会用例子来说明这一技术，另外在第 12 章讨论高速无线局域网（WLAN）时也会对其进行简要的说明。垂直频分复用（OFDM）被用来平行地发送数字信号序列，减小所需带宽，从而使信道产生平衰落，随之减少了符号间的干扰，这一技术将在第 5 章中讨论，另外在第 12 章讲述高速 WLAN 时将对其应用进行说明。编码技术也被广泛地用于检测或纠正在信号传输过程中发生的数字信号错误；编码过程将在第 7 章中讨论。

在本节中，讨论 3 个其他的用于改善信号衰落影响的技术。第 1 个是使用均衡器，它被设计用来克服或者至少减少符号间干扰的；第 2 个是分集技术，对给定传输或接收信号进行多路复制，接收机对其进行合适的组合可以提高系统的性能；第 3 个是 RAKE 接收机，它提供了分集接收的一种特殊情形，适合用于使用很宽带宽的无线系统，像使用 CDMA（码分复用）技术的系统，在后面章节所要讨论的第二代和第三代系统中都有使用。这些技术在移动无线系统中都有很长的使用历史。早在 1927 年，分集技术就可能被使用了，用来克服高频信号传输包括电离层的反射信号遇到的短期衰落，从那以后很快地在 1931 年 RCA 在无线接收系统中并入了分集技术（见 Schwartz 等，1996）。在本节中要讲述的一些分集方案相对成熟的分析在 20 世纪 50 年代已完成（Brenner，1959）[①]。RAKE 接收机作为对抗多径衰落效应的方法被引入也是

---

① Brenner, D. G. 1959. "Linear diversity combining techniques," *Proc, IRE*, 47, 6(June), 1075 – 1102.

在20世纪50年代(Price和Green,1958)①。均衡器在20世纪60年代中期被首次开发,并在当时用做电话的调制解调器,以克服数字电话传输中的符号间干扰(Lucky,1965)②。

均衡技术

如上所述,均衡器的目的是用来消除或者至少是减小符号间的干扰的。要消除信道中像无线信道的符号间干扰,需要对信道进行估计和模拟。均衡是在基带上进行的,接收到的失真信号首先被解调到可携带信息的频率范围。下面所有的讨论都只考虑基带情况下,这包括对信道的估计和模拟。现在考虑对调制载波信号传输实际信道的基带等效。对信道的估计通常是通过发射指定的数字测试序列,在知道序列的情况下调整均衡器来使一些想要的性能目标最小化。图2.20给出了一般使用的均衡器的形式,此设备由一个线性不回归或者说横向滤波器构成,如图2.20所示,有$2N+1$个抽头。假设在一个符号间隔发射的数字信号通过信道被扩展成长为$2N+1$个符号间隔,图2.21(a)所示的数字信号可以作为一个例子。图中在一个符号间隔发射的信号在接收端将在某段时间后达到最高点,这考虑到信道传播时间及其之后的延迟。第2个例子见图2.21(b),接收信号在时间上滞后。这两种情况都是在每个符号间隔的中间进行采样,得到均衡器的输入,这些采样值在图2.21中显示出来。实际中,输入采样序列作为均衡器的输入,这些输入采样值每隔T秒符号间隔更新一次,第j个这样间隔的采样值被记为r_j。从图2.20可看到,这些采样值要通过$2N+1$个T s的延时单元,每一个延时单元的输出都将乘上一个抽头系数,然后这些结果的和用来形成对发射基带信号s_j的估计$\hat{s}_j$。对应于一个符号间隔j接收到的采样值,均衡器的输出为

$$\hat{s}_j = \sum_{n=-N}^{N} h_n r_{j-n} \tag{2.51}$$

注意到输出在形式上正好是一个数字卷积,类似于前面在式(2.30)~式(2.32)中所讨论的卷积,而这也在预料之中,因此图2.20的设备是一个数字滤波器。

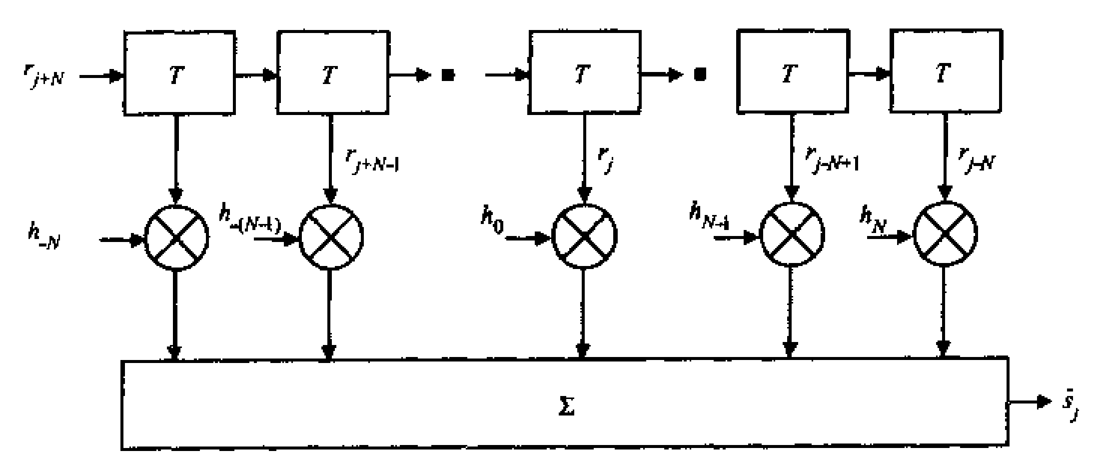


图2.20 横向滤波均衡器

现在的问题是要确定$2N+1$个抽头系数h_n,$-N \leq n \leq N$。一个常用的技术就是先发射一个已知的长为K比特的训练序列,在第6章中将详细讨论为此使用的伪随机序列。在接收

① Price, R. and P.E. Green, Jr. 1958. "A communication technique for multipath channels," Proc. IRE, 46, 555-570.

② Lucky, R.W. 1965. "Automatic equalization for digital communication," *Bell System Technical Journal*, 44 (April), 558-570.

端发射序列和估计序列都已知，可对抽头系数进行调整使发射信号的值和估计信号的值之间的差别最小。由于信道特性随时间变化，所以均衡器的抽头增益也应该做相应的调整，可利用发射的训练序列来实现。这里假设增益的调整与信道特征的变化相比能很快地完成，信道的变化很大程度上是由无线终端的移动造成的。依照长为 K 比特的训练序列，一个用来确定均衡器抽头增益的简单性能指标，是使发射信号序列 s_j 与估计信号序列 $\hat{s}_j$ 之间差值的均方值（$1\leqslant j\leqslant K$）达到最小。

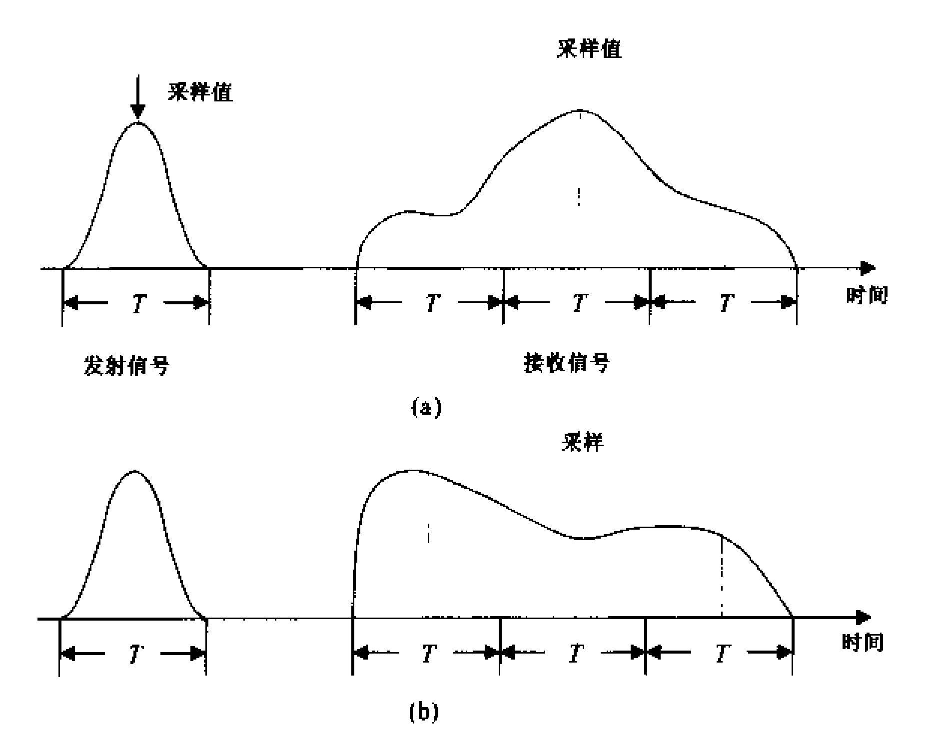


图 2.21　信道响应的例子

最小均方指标

$$\text{找到 } h_n,\ -N\leqslant n\leqslant N,\qquad \text{使得 } \sum_{j=1}^{K}(s_j-\hat{s}_j)^2 \text{ 最小} \tag{2.52}$$

式(2.52)所要求的最小值很容易得到，关于 $2N+1$ 个滤波器系数 h_n，利用式(2.51)把式(2.52)中的 $\hat{s}_j$ 替换为接收信号采样值 r_j，并令每一项都为零，则可得到 $2N+1$ 个等式

$$\sum_{j=1}^{K}s_j r_{j-l}=\sum_{j=1}^{K}\sum_{n=-N}^{N}h_n r_{j-n}r_{j-l}\qquad -N\leqslant l\leqslant N \tag{2.53}$$

这 $2N+1$ 个等式必须同步解出来，从而得到想要的抽头系数 h_n，$-N\leqslant n\leqslant N$。这一最小化过程看起来很可怕，不过通过定义下面两个系数可以使之看起来要好很多：

$$R_{l,n}\equiv\sum_{j=1}^{K}r_{j-n}r_{j-l}\qquad -N\leqslant l,n\leqslant N \tag{2.54}$$

$$g_l\equiv\sum_{j=1}^{K}s_j r_{j-l}\qquad -N\leqslant l\leqslant N \tag{2.55}$$

式(2.53)可简化其形式为

$$g_l = \sum_{n=-N}^{N} h_n R_{l,n} \qquad -N \leqslant l \leqslant N \tag{2.53a}$$

由式(2.54)定义的系数 $R_{l,n}$ 称为采样相关系数，而且事实上它与前面由式(2.35)定义并讨论的自相关函数有着紧密的联系。采样系数事实上是在实践中确定自相关函数的一种方法，可通过替换式(2.35)的总体平均为式(2.54)的采样均值来实现(Schwartz 和 Shaw,1975)①。这里需要假设自相关所要估计的系统在测量时间内是不变的，而且 K 个采样值在数量上足够多，足以对自相关进行不错的估计(Schwartz 和 Shaw,1975)。

式(2.53a)的 $2N+1$ 个等式可易写为向量形式，进一步简化求解均衡器系数的分析形式。定义 $2N+1$ 元的向量 $\boldsymbol{g}$ 和 $\boldsymbol{h}$，其元素分别为 g_l 和 h_n，再定义一个 $2N+1 \times 2N+1$ 的矩阵 $\boldsymbol{R}$，其元素为 $R_{l,n}$，易得式(2.53a)的向量形式为

$$\boldsymbol{g} = \boldsymbol{Rh} \tag{2.56}$$

$2N+1$ 个滤波器系数值的向量解法为

$$\boldsymbol{h} = \boldsymbol{R}^{-1}\boldsymbol{g} \tag{2.57}$$

如果采样数 K 很大并且作为训练序列的伪随机序列其正负二进制数字的数量差不多，则对式(2.53)、式(2.56)和式(2.57)的 $2N+1$ 个等式的同步解可以简化，令采样值 s_j 是 $+1$ 或者 -1。我们将在第 6 章中更进一步地讨论这些序列，这些相继的比特之间差不多相互独立，进一步假设在传输过程中发生的错误比特与比特之间是随机独立的，则相继接收到的比特之间也是相互独立的。在这些条件下，可以看到由式(2.54)定义的 $R_{l,n}$ 在 $l \neq n$ 时近似为零，在 $l = n$ 时为 K，则矩阵 $\boldsymbol{R}$ 接近为对角阵。对滤波器系数一个好的近似如下：

$$h_n \approx g_n/K = \sum_{j=1}^{K} s_j r_{j-n}/K \qquad -N \leqslant n \leqslant N \tag{2.58}$$

在求图 2.20 所示的横向滤波器的抽头系数时，上面所用的最小均方准则与对这些系数求最大似然估计等效(Schwartz 和 Shaw,1975)。最大似然估计在第 7 章中用于卷积码的维特比译码，Steele(1992)描述了一种基于维特比译码并为 GSM 蜂窝系统设计的均衡器。

除了最小均方指标或者等效的最大似然估计之外，其他的标准也可用于设计均衡器以消除或减小符号间干扰。由 Mark 和 Zhuang(2003)② 详细讨论的迫零均衡器就是一个这样的设备，由其名称可以看出，这种均衡器的输出 $\hat{s}_j$ 在第 j 个采样值与发射信号 s_j 相等，即令信道产生的符号间干扰采样值对 s_j 的响应为零。在线性 z 变换域，这一过程相当于令总的传输函数 $H(z)$，即信道传输函数和均衡器传输函数的乘积为 1。如 Mark 和 Zhuang(2003)所述，这一过程相当于均衡系统的冲激响应在符号采样时刻为 1，而在其他符号采样时刻为 0。这一结果当均衡器跨度为 $2N+1$ 个符号间隔时是可能存在的，像在最小均方均衡器中需要解 $2N+1$ 个方程一样。其他类型的均衡器，Mark 和 Zhuang(2003)也都有叙述。

分集接收

本节要讨论的改善数字信号传输中衰落效应的第 2 种方法就是分集接收：数字信号多重独

① Schwartz, M. and L.Shaw. 1975. *Signal Processing: Discrete Spectral Analysis, Detection, and Estimation*, New York, McGraw-Hill.

② Mark, J. W. and W. Zhuang. 2003. *Wireless Communications and Networking*, Upper Saddle River, NJ, Pearson Education Inc.

立的样值在每个符号间隔接收或者发射,用来降低此间隔发射信号的检测错误概率(Schwartz等,1966)。这一技术在无线传输中使用历史很长,是利用多个统计样值来改善被污染信号的检测和估计性能的一个例子。多信号样值可由多天线来获得,这一技术常称为空间分集。若由多频信道传输得到的话,则称为频率分集;若在时间上对给定信号接收多个样值的话,则称为时间分集。分集技术的其他例子还包括角度分集和极化分集,在前一种情况下,需使用指向不同方向的方向性接收天线,从而在这些方向接收到独立散射的信号;后一种情况,信号需要在两个垂直极化的方向上同步发射。多天线的使用,可以在发射端,或在接收端,也可以两端都有。回想在瑞利衰落情况下,衰落特性在以波长为阶的距离上发生改变,于是间距为半个波长的天线就可提供接近独立的信号样值。一个更普遍的用于克服衰落和其他信号干扰的技术术语为 MIMO(多入多出)接收,在发射机和接收机都使用了多天线,由于它在信号接收上可提供大改进的潜力,近几年这一技术很热门,在第 6 章中给出了有关这一技术的简要介绍和参考书目。

考虑一般形式的分集接收,假设在每一个符号间隔可得到 N 个独立分集样值,不管使用的是哪种分集技术,都可以模型化为传输信号经过 N 个独立的瑞利衰落信道,每个信道的基带输出可作为在每一个间隔的分集样值。将 N 个瑞利信道的输出线性合并来提高分集性能,线性合并器输出的判决用来确定每个符号间隔发送的数字信号。用分集技术来改进无线通信系统的过程如图 2.22 所示。用式(2.30)的形式来表示瑞利信道的输出基带信号,用 $a_k(t)$代替 $a(t)$来表示第 k 个分集信道。为简单起见,假设要恢复的基带信号是二进制的,为 $\pm s(t)$、在对 N 个分集信道信号进行恰当的合并后,判决器将在每一个二进制间隔进行判决,即决定所传输信号是正还是负。

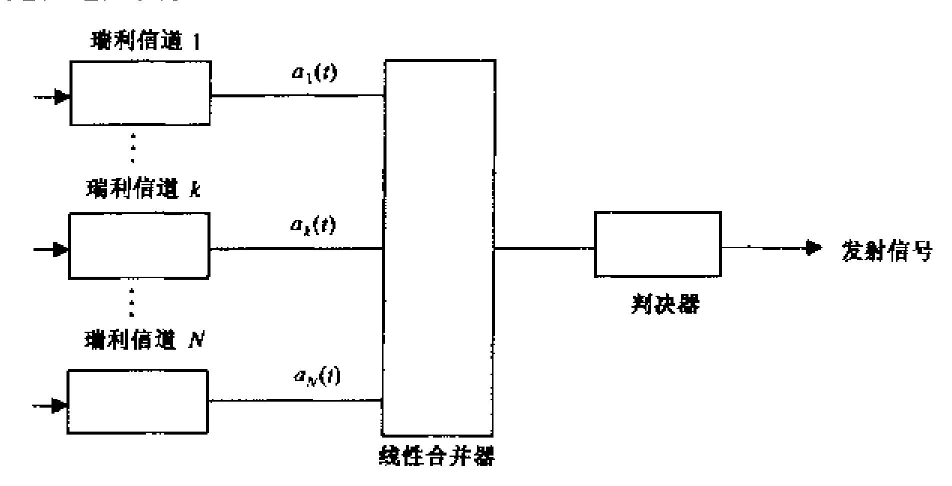


图 2.22　对 N 路分集线性合并的模型

进行线性合并 N 路分集信号有 3 种常用方法(Brenner,1959;Schwartz 等,1966):选择性合并、等增益合并和最大比合并。选择性合并是最简单的操作,合并器在 N 路分集信道中选择一路有最大功率的作为判决信号①。等增益合并是将 N 个信道的输出在移除相位差异后加起来。最大比合并是这 3 种技术中最复杂的,它在将每一个信道的输出相加之前,考虑它们的幅度并对其进

① 许多作者使用第 k 路分集支路的信噪比 SNR 作为合并器的输入。下节将要引入的概念信干比 SIR,在衰落环境下经常替代 SNR 使用,尤其是当干扰在无线信道中占主导作用的情况下。不过,由于每一个分集支路上的"噪声功率"或者"干扰功率"的模型是一样的,所以不论是用 SNR 还是 SIR,或只是功率,在我们的分析中都不会有什么差别。为了简单起见,这里我们只关注功率,在接下来进行的最大比合并的讨论中,再考虑干扰和噪声。

行权重，同样此时也要将相位差异除去。本节后面将看到最大比合并是这 3 种合并技术中性能最好的，不过，在小 N 分集情况下，相对于选择性合并，其性能提高不大。可以通过很多方法来测量性能，在后面两章以及本书的其他地方都主要将信干比 SIR 作为移动无线系统性能的量度。在后面章节中可以看到，一个性能好的系统就是其 SIR 超过一个特定阈值的系统。由于选择合并对分集带来的系统性能的提高也将从这一参数得到体现，SIR 也决定了干扰有限无线系统的误比特性能（干扰有限意味着，对于系统性能的影响来说，相对总存在的噪声干扰起着主导的作用）。将在第 6 章的 6.4 节讨论由于双分集（例如，两根天线）而带来的误比特性能的改善。

现在我们来看选择性合并——分集合并技术中最简单的一种方法，可看到即使有限数量的分集（小 N）也可以显著改善系统性能。回想式(2.23)，对于瑞利信道瞬时功率服从均值为局部平均功率 p 的指数分布，为简单起见，假定 N 个瑞利信道的局部平均功率一样（见 Schwartz 等，1966，在不等的局部平均功率条件下的选择性合并：不过，那里的讨论使用的是 SNR，而不是功率）。由式(2.23)可知，瞬时功率的概率分布或小于某一阈值 γ 的瑞利信道的瞬时功率概率为 $(1-e^{-\gamma/p})$。现在考虑 N 个这样的信道，N 个独立瑞利信道的瞬时功率小于 γ 的概率为这些概率的乘积，而且至少有一个信道的瞬时功率大于 γ 的概率为

$$\Pr[\text{最少有一个信道 } P>\gamma]=1-(1-e^{-\gamma/p})^{N} \tag{2.59}$$

所有信道瞬时功率都小于此阈值的概率，有时称为中断率（Schwartz 等，1966）

$$\Pr[\text{所有信道} Ps<\gamma]=(1-e^{-\gamma/p})^{N} \tag{2.60}$$

由式(2.60)可以看出，随着分集的增大，中断率下降得很快。相反地，至少有一个瞬时功率超过阈值的概率，无线系统的一个良好的品质，随着分集的增大将很快地到达 1，因此有限数量的分集也能在克服短期衰落方面使性能得到显著提高。考虑下面一些例子：

例 1

使用双选择分集，将它的阈值性能与没有分集的相比较，阈值为 10(10 dB)，局部平均功率是 100(20 dB)。这里，$\gamma/p=0.1$，则 $e^{-0.1}=0.905$，这是一个信道超过阈值的概率，于是信道功率在阈值以下的概率为 0.095。两个信道中至少有一个信道功率超过阈值的概率为 $1-(0.095)^2=0.991$。双分集选择性合并输出低于阈值的概率为 $0.095^2\approx0.009$，这与没有分集的单信道情况相比有数量级的减小。每加上一路分集都将使低于阈值的概率以 $0.095\approx0.1$ 的倍数减小。

（注意，dB 或分贝是和功率比相关的，分贝的定义为

$$\text{dB 数}=10\lg(\text{功率比})$$

上面的数字隐含着与功率 1 相比的那些功率值。功率比为 2 等价于 3 dB，功率比为 10 等价于 10 dB，等等）。

例 2

考虑局部平均功率与阈值可比的情况。特别地，假设 $\gamma/p=1$，在不使用分集的情况下，瞬时功率超过阈值的概率为 $e^{-1}\approx0.368$。中断率，功率降到阈值以下的概率为 0.632，这是一个相对较高的值，于是此时系统工作状态不好。双分集降低了这一概率，合并其输出功率，则在阈值以下的概率为 $0.368^2=0.135$。4 分集系统输出功率在阈值之下的概率降为 $0.135^2=0.018$，4 路中至少 1 路的瞬时功率大于阈值的概率为 0.98，这是一个相当好的结果。

现在来看最大比合并，这里为把合并器的输入信号的形式普遍化，把每一路 N 瑞利信道的干扰和噪声看做是加性的。第 k 个信道的输出为 $a_k(t)+n_k(t)$，其中 $n_k(t)$ 表示干扰和噪

声。作为合并的一部分,第 k 个分集信道的输出与一个增益因子 g_k 相乘,所有信号的输入彼此没有相位差,然后信号与干扰的和将应用到图 2.22 的判决器中。信号与干扰的和在每一个间隔的采样值将用于确定发射的数字信号[回想式(2.30)中复信号 $a_k(t)$ 的幅度为 $|a_k(t)|$,相位为 $\theta_k(t)$,对于最大比合并,假定调制器可以估计每一个角度 $\theta_k(t)$,并将它们除去,所以所有信号之间没有相位差]。于是,在采样时刻最大比合并器的输出有两个分量,信号部分为 $s=\sum_{k=1}^{N} g_k a_k$,其中 a_k 为 $a_k(t)$ 在采样时刻的值,干扰加噪声的部分为 $\sum_{k=1}^{N} g_k n'_k$,其中 n'_k 为合并器输出的第 k 路分集信道干扰/噪声项的采样值。现在的问题是怎么选择增益因子 g_k,$1\leqslant k\leqslant N$。最大比合并中采用的方法是选择增益因子,使得采样时刻合并器输出的信干功率比 SIR 最大(再次注意,许多作者使用的是信噪功率比 SNR,这里选择用 SIR 主要是因为在移动无线系统中干扰是最常遇到的有害量。不过,在这个讨论中,干扰和噪声都包括在内)。合并器输出的信号功率与 $|s|^2$ 成正比,为确定干扰加噪声的功率,假定每一路的干扰彼此独立,于是在合并器的输出端单个的干扰和噪声功率相加,得到 $I_o=\sum_{k=1}^{N}|g_k^2||n'_k|^2$。采样时刻的 SIR 与比值 $|s|^2/I_o$ 成比例。于是,最大比合并便成了选择单独的增益 g_k 来最大化 SIR:

最大比合并

对所有的 k,选择 g_k,使得下式最大

$$SIR\equiv|s|^2\Big/I_o=\left|\sum_{k=1}^{N} g_k a_k\right|^2\Big/\sum_{k=1}^{N}|g_k|^2|n'_k|^2 \tag{2.61}$$

应用复数的施瓦兹不等式可使该比值最大化,施瓦兹不等式为

$$\left|\sum_{k=1}^{N} c_k^* d_k\right|^2\leqslant\left[\sum_{k=1}^{N}|c_k|^2\right]\left[\sum_{k=1}^{N}|d_k|^2\right] \tag{2.62}$$

当 $d_k=Kc_k$ 时,等式成立,K 是一任意常数。星号 * 表示复共轭运算。

为了将式(2.61)写成式(2.62)可应用的形式,令 $c_k=a_k^*/|n'_k|$,$d_k=g_k|n'_k|$,于是由式(2.62)(Schwartz 等,1966)可得

$$SIR\leqslant\sum_{k=1}^{N}|a_k|^2/|n'_k|^2 \tag{2.63}$$

对于第 k 路分集,$k=1$ 到 N,当

$$g_k=Ka_k^*/|n'_k|^2 \tag{2.64}$$

时,等号成立。由于 $|a_k|^2/|n'_k|^2$ 只是第 k 路分集的 SIR,则结果就等于这些 SIR 的和。简单而直观的解释如下:信号幅度相对干扰/噪声功率大的支路选择高的增益因子;对于信号幅度相对干扰/噪声功率小的支路减小其增益因子。这种类型的最优化称为匹配滤波器,在通信理论中,当在有噪声的情况下检测信号时会用到这种滤波器。在滤波器的输出端使信噪比最大,这种意义的滤波器就是匹配滤波器。更一般地,这个最优接收系统是相关检测器的一个例子,在相关检测器中要检测的是实际发射的多路信号中的一路(Schwartz,1990)。注意到通信理论情况下的问题与这里的问题类似,通信理论中最优化问题和匹配滤波器方案在雷达信号检测中首次得到了发展。这种滤波器的另一个解释是在频域方面,当频带中信号功率大于噪声时,提高信号功率;当频带中噪声大于信号功率时,降低信号功率(Schwartz,1990)(在这里,更精确的描述应该是信号和噪声的功率谱密度)。

最大比合并、等增益合并(所有分集增益相同,因此,分集信道的输出至少简单地相加,而不是对每一路用增益因子进行权重)和选择性合并之间的性能比较已经被人们研究过(Schwartz 等,1966)。有趣的是等增益合并的性能几乎与最大比合并一样好,不过当分集信道数目增加时,性能差异会增加,但对于 8 路分集,这一差异仍小于 2 dB。在小数量分集时,选择性合并的性能也相当接近最大比合并:对于两路分集性能差异大约为 1 dB。当合并信道数目增多时,性能差异增大:3 路时为 2 dB,4 路时稍小于 4 dB,8 路时为 6 dB。

如前所述,最大比合并需要知道每一路分集信道的相位和幅度信息。关于这两类信息变化偏差的研究表明,当相位偏差为 37.5°时,性能恶化不超过 1 dB;当信号幅度偏差为 0.5 dB 时,性能偏差也不超过 1 dB(Schwartz 等,1966)。

RAKE 接收机

在本节的引言中已经提到,RAKE 接收机对具有较宽带宽的无线系统如码分复用 CDMA 将使其性能得到很大的改善。考虑一个工作在高符号速率的数字无线系统,其速率足够高以至于衰落信道的时延扩展大于符号周期,这一情况如式(2.48)所描述,将导致频率选择性衰落。随着符号间隔减小,信号带宽加大,多径信号的单个分量到达时分别有不同的时延,可能会被分离开。如果单个多径分量的相对时延同相对相位和相对幅度一样可以被准确地估计,那么不同的接收波束可以在时间上进行平移来补偿相对时延。最大比合并可用于时间补偿的到达波束。RAKE 接收机方案是将通过衰落信道的发射信号的独立到达波束进行合并,在时间补偿和合并后,波束被分离开,这与耙子或叉子的爪很像,所以称为 RAKE 接收机。图 2.22 提供了这种“耙状”结构,前面已提过,这一时间分集的技术首先被 Price 和 Green(1958)所论述(详细的论述见 Schwartz 等,1966;Rappaport,2002)。

在本书后面所提的 CDMA 系统中,每一个单个的二进制数字都将乘以一个更高比特率的伪随机二进制序列,从而得到非常小的符号间隔,称为 CDMA“码片”。这样的系统也称为扩频系统,因为乘以伪随机序列后传输带宽增加了,也就是说扩展了频谱。该系统的用户被分别分配一个不同的伪随机序列,这一序列对发射机和接收机都是已知的,也就是 CDMA 系统中的“码”。然后,编码序列对载波进行调制,并且通过无线信道传输。假设信道有多径衰落且时延扩展大于码片间隔,那么在接收端单个的波束将被分离开。接收系统可以对接收到波束的不同时延值进行很快的扫描,找到对应于发射波的序列。一旦一定数量的这种有相对时延的接收信号被鉴别出来,就可通过时延补偿和最大比合并来恢复发射序列。在实际中,最早到达的波束很少,通常功率上最强的波束用于 RAKE 接收。在第 6 章中将给出一个实例,通过使用 RAKE 接收机来降低 CDMA 系统中误比特性能。

习题

2.1 用自己的话解释一些平均路径损耗、阴影衰落和多径衰落之间的差别,它们之间是怎么互相联系的?

2.2 利用表 2.1,画出奥兰多的 $P_{R,dB}-P_{T,dB}$ 曲线,以距离 d 为自变量,d 的单位为 m,且 $0<d<200$ m。假定发射天线和接收天线的增益均为 1。

2.3 分别对图2.4和图2.5中4条测量曲线求其阴影衰落参数 σ，并比较。提示：首先计算出每条曲线的平均值，然后计算这些均值的均方根。

2.4 距离基站100 m的移动台接收到的平均功率为1 mW，这段距离中经历了对数正态阴影衰落

(a) 距离基站100 m的移动台接收到的功率大于1 mW的概率是多少？小于1 mW的概率是多少？

(b) 对数正态标准差 σ 是6 dB，可接受的接收信号功率应为10 mW或者更高。移动台可接受信号的概率有多大？当 σ 为10 dB时呢？如果可接受的接收信号的功率为6 mW的话，对 σ 的两个值再分别求一下移动台可接受的接收信号的概率。注意，这里的解答需要高斯函数的积分，大多数数学软件都有这种方法。大多数概率统计的书都有用于此目的差错函数的表格，在第3章中给出有差错函数的定义见式(3.12)。

2.5 (a) 详细写出如式(2.13a)所示两波束平均接收功率的推导过程。

(b) 在图2.7和图2.8中添加一个 $1/d^4$ 的曲线，并与其他测量曲线相比较。

(c) 图2.8的结果是否验证了式(2.13a)？并解释。

2.6 (a) 证明式(2.15)中对于瑞利分布随机变量 a，σ_r^2 必须等于1/2。

(b) 由式(2.17)推导出式(2.17a)，证明 x 和 y 都是均值为零，方差为 σ_R^2 的随机变量。

(c) 利用接收信号的复包络 α 的瑞利分布式(2.21)，证明瞬时接收功率 P_R 服从式(2.23)所示的指数分布。

2.7 (a) 由文中所给的提示，证明莱斯分布式(2.25)在 $A^2/2\sigma_R^2 \gg 1$ 情况下逼近一个以 A 为中心的高斯分布。

(b) 证明在直接到达波束情况下，瞬时接收功率分布为式(2.27)。证明随着因子 K 减小(视距直接波束相对散射信号项减少)，衰落分布式(2.27)逼近于指数分布。

2.8 从文中可以看到，仿真与证明分析一样经常用来确定蜂窝系统的性能。这里最重要的是衰落环境的仿真，在此提供了一个对瑞利衰落仿真的介绍。

(a) 考虑一个有 n 个随机数 x_j 的序列，$j=1$ 到 n，在0到1之间均匀分布(在数学软件中伪随机数字发生器很常见)。令 $x=(b/n)\sum_{j=1}^{n}(x_j-1/2)$，证明 x 近似为一个均值为0，方差 $\sigma^2=b^2/12n$ 的高斯随机变量。考虑另外 n 个均匀分布的随机数(独立的)，这时其和为 y，用 x 和 y 生成一个瑞利分布随机变量。与式(2.21)比较，这时瑞利参数 σ_R^2 是什么？提示：考虑从式(2.19)到式(2.21)的推导和下面文章的讨论。

(b) 另外一个得到瑞利分布的方法是直接由一个均匀分布的随机变量 x 得到的

$$a=\sqrt{-2\sigma^2\ln x}$$

证明变量 a 服从瑞利分布。你如何用一个均匀分布随机数序列来生成瑞利分布？

(c) 对 n 选择不同的值，用(a)和(b)两种方法来生成瑞利分布，比较它们，并与瑞利分布图相比较。n 的变化将对这些结果产生怎样的影响？

(d) 证明你怎样通过扩展(a)的方法来得到莱斯分布，对莱斯 K 因子的不同值，用(c)中生成的伪随机数来得到，并画出这一分布。

2.9 考虑平均衰落持续时间等式(2.39)。如果车辆行驶速度为 100 km/h，系统工作频率为 1 GHz，假定比率 $\rho=1$，证明平均衰落持续时间为 8 ms，如文中所述。现在令接收信号幅度为其均方根值的 0.3，证明平均衰落持续时间为 1 ms。

2.10 (a) 用自己的话对文中关于时间色散和频率选择性衰落的讨论进行概括。

(b) 考虑几种情况：时延扩展为 0.5 μs，1 μs，6 μs，对于这两个传输带宽，用于 IS-95 和 cdma2000中的 1.25 MHz 和用于 WCDMA 中的 5 MHz，单个的多径波束是否能被分离开(见第 10 章)？

2.11 指出对于下列每个数据速率来说发生平衰落的条件

8 Kbps，40 Kbps，100 Kbps，6 Kbps

并指出对于每一个数据速率发生平衰落的无线传输环境(如果有的话)。

2.12 (a) 考虑图 2.20 所示的横向滤波均衡器，一个有 K 个二进制数字的训练序列被用于确定 $2N+1$ 个抽头系数。证明在最小均方性能目的下，抽头系数的最优选择如式(2.53)所示。

(b) 证明式(2.53)的向量形式式(2.56)，其解答为式(2.57)。

2.13 (a) 设计出一个横向滤波均衡器的简单例子：假定均衡器有 3 个抽头，使用最小均方性能准则，任意选择 $K=10$ 个二进制数字作为训练序列，并令其中一些为接收“错误”，也就是说其中一些变为相反的极性。试着选择训练序列使其中的 +1 和 -1 数目相等。将这些抽头系数与通过式(2.58)得到的相比较。

(b) 用另外一套不同的发射数字信号和接收错误来重新做这个例子。

(c) 选择一个大型的横向滤波均衡器来重做(a)和(b)，并比较得到的结果。

2.14 (a) 绘出使用双选择分集时获得的性能提高随局部平均功率和阈值比值变化的曲线。至少考虑下列几种情况：(1)局部平均功率为阈值的 20 倍；(2) 局部平均功率为阈值的 10 倍；(3) 局部平均功率与阈值相等；(4) 局部平均功率为阈值的 0.1。注意，性能可被定义为中断率或者至少有一个瞬时功率大于阈值的概率。

(b) 对于 4 路分集和 8 路分集重复(a)，并比较这 3 种分集结果。

2.15 证明对于第 k 路分集，最佳最大比合并增益为式(2.64)。解释 SIR 是 N 个分集支路 SIR 的和。

2.16 (a) 解释等增益合并与最大比合并的不同。特别地，写出等增益合并的 SIR 表达式。提示：这一表达式与式(2.61)比较会如何？

(b) 为什么你会认为分集方案的性能排行顺序为：最大比合并，等增益合并，选择性合并？

2.17 (a) 用自己的话解释 RAKE 接收机的工作原理。

(b) 在第 10 章中讨论了两个第三代 CDMA 系统，一个是 WCDMA，码片速率为 3.84 Mcps(百万码片每秒)，相应的码片持续时间为 0.26 μs；另一个是 cdma2000，码片速率为 1.2288 Mcps，码片持续时间为 0.81 μs(这与第 6 章和第 8 章讨论的第二代 CDMA 系统 IS-95 的码片速率一样)。解释第 10 章所说的，分别对于这两个系统，RAKE 接收机可用来提供多径时间分集，在至少两个码片持续时间上有不同的路径。哪个系统能提供更好的 RAKE 性能？

第3章　小区基本概念和信道分配

3.1　信道再用和小区导论

在前面一章里，我们已经介绍了在基站和移动终端之间由于多径传播造成的影响。其中注意到，大部分现代无线系统被规划为许多小区，每个小区由一个基站来控制（也有一些例外，一些小范围地区系统，比如本地无线网和个人无线网，这将在第12章进行讨论。本书不涉及*Ad hoc*和无线传感器网络，因为它们属于其他类型的网络，跟小区概念没有关系）。我们也将讨论一些在单小区范围内的传播影响。本章介绍了小区基本概念和怎样使用小区才能增加系统容量，使得更多的用户可以同时进行通信。

一个无线系统可容纳同时呼叫的数量主要由以下因素决定：系统频谱如何分配，一个呼叫信号所占用的带宽。这与其他的系统并没有本质的区别，如电台广播系统AM或FM，电视广播系统等。举个例子，美国的第一代模拟无线通信系统AMPS，使用800～900 MHz无线频段，上下行各自占用25 MHz频谱。在无线频谱分配方面，联合国通信委员会（FCC）在1981年分配824～849 MHz频带给上行链路（也称反向链路，从移动台到基站）；869～894 MHz频带给下行链路（也称前向链路，基站到移动台）。第一代采用频率调试方式（FM），上下行各25 MHz带宽，每个呼叫使用30 kHz带宽，所以每个方向有832个信道可以使用。第二代和第三代数字无线通信系统使用时分多址（TDMA）和码分多址（CDMA）技术，在同样25 MHz带宽的情况下可以使得系统容量进一步上升，容纳更多的用户同时进行呼叫。时分多址和码分多址的概念，以及第二代系统和第三代系统的具体描述将在以后的章节中进行讨论，但是它们最终都要受限于能分配多少频谱。第一代832个模拟信道以及使用数字技术后翻了几倍之后的信道数量，仍然不能满足巨大的用户数量，尤其是城市和郊区。把一个地区分割成许多地理上不同的小区，每个小区称为一个蜂窝，可以进行频率再用，这样可使通信信道的数量明显增加。小区的概念就是将一个地区分隔成几个地理上不同的小的区域，每个由一个基站控制。这种方法最早由美国的贝尔实验室于1947年首先提出。高容量无线电话系统与小区概念结合的具体方案在1971年由贝尔实验室提交给FCC（Jakes，1974）[①]。第1个AMPS系统于1983年首先在芝加哥获得使用。

如前所述，小区的基本思想就是在不同的小区重新使用频率来增加系统的容量。但是这样会有个新问题：同样的频率不能在临近的小区分配，否则会造成信道间的干扰。使用同样频率的小区必须保持足够远，直至达到系统能忍受的干扰水平。即使这样，比起预期的目标，信道还是不够用。但是只要小区数量足够多，就可以提供更大的系统容量。如果小区的尺寸可以减小，那么更多的小区可以被增加到一个特定的地理区域中，这样也可以增加系统的容量。近来，使用微蜂窝技术是一个增加系统容量的趋势。

通过小区的引人，我们已经解释了小区容量是如何改善的。考虑一个简单的一维小区的

---

① Jakes, W.C. ed. 1974. *Microwave Mobile Communications*, AT&T, 1995 edition, NeW York, IEEE Press.

例子(这种模型可以用来建模公路上小区的分布情况)。首先是第一代 AMPS 系统,总共 832 个信道被分成 4 组,每组 208 个信道。我们称这些组为 1,2,3,4,然后把它们放在如图 3.1 所示的位置。这样,两个使用同样频率的小区之间隔着 3 个小区,称为 4 小区再用。给定小区数量 N,总共有 $208N$ 个信道可以使用,可见与原先没有小区时的信道数量相比有了很大的提高。如果小区数量 N 足够大,那么就可以使系统容量比原来的 832 个信道得到一个很大的提高。如果仅仅使用两个小区来分隔具有同样频率的小区(即 3 小区再用),将得到 $277N$ 个有效信道,使得系统容量得到更大的改善。那么,应该使用哪种策略呢?这主要依靠系统抗干扰的能力,因为 3 小区再用拥有同样频率的小区间隔比较小,而 4 小区再用相对间隔大一些,所以干扰相对也就小一些。这种可容忍的干扰主要由系统的信干比(SIR)决定,也称为载干比(CIR)。这个比例定义为在接收机端接收到的平均信号功率与平均干扰功率之比。对比于信噪比(SNR)有所不同,其主要在非移动通信系统中使用(Schwartz,1990)[1]。如果想正确接收信号,SIR 应该大于一个指定的阈值。例如,第一代 AMPS 系统主要为语音通信而设计,期望的性能对应阈值为18 dB($SIN_{dB} = 10\lg SIR$)。在研究移动通信呼叫方面,发现这个阈值可以提供比较可靠的声音质量。第二代数字 AMPS 系统(D-AMPS 和 IS-54/136,将在以后章节中讨论),它们的阈值可以设定在 14 dB。GSM 系统(也将在后面进行讨论)已经在世界上得到了广泛的使用,通过研究发现,它的 SIR 范围在 7~12 dB 比较合适。我们也将在以后的章节中看到,数字通信系统的错误概率主要依靠这个阈值的设定。

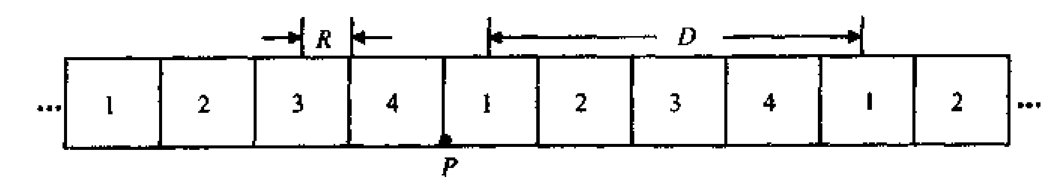


图 3.1 一维小区阵列示意图

3.2 信干比计算,一维情况

以上强调信干比计算主要依据平均功率,然而应该注意到,信干比的计算主要由式(2.4)中提到的距离决定。阴影衰落和多径瑞利衰落可以被忽略。尤其应该注意到假设式(2.4)中的 $g(d)$ 为 kd^{-n} 形式,其中 n 一般取值为 3 或者 4,后者使用一个理想的两径模型得到,见式(2.13a)。首先计算一维情况的信干比,然后考虑二维的情况。在一维小区的例子中,D 是干扰小区之间的间隔,小区一半的宽度(即中心到每个边的距离)为 R,如图 3.1 所示。观察移动台接收的下行功率,最差的地方是在小区的边沿。这个位置在图 3.1 中用点 P 指明(同样的计算方法可以应用于上行链路,实际上,在研究 CDMA 系统的时候就已经用到了这种方法)。比如,每个基站在小区的中央,发射同样的功率 P_T(这意味着功率可以被控制,关于这点将在以后叙述)。距离基站 d 米远的平均接收功率是 $P_T d^{-n}$,n 为 3 或 4,在点 P 的移动台的接收信干比为

$$SIR = \frac{R^{-n}}{\sum P_{int}} \tag{3.1}$$

---

[1] Schwartz, M. 1990. *Information Transmission, Modulation, and Noise*, 4th edn, New York McGraw-Hill.

其中,分母包括所有干扰的基站,P_{int}表示归一化的干扰功率。注意,计算归一化平均功率是因为先假定所有的发射信号功率都是一样的。可以清晰地看到,至少在理论上,发送同样频率的其他基站信号会干扰主基站的信号发送。通过例子表明,由于随着距离 d 的增加,接收信号能量将迅速减小,所以仅仅需要考虑邻近的一些干扰。

当仅仅考虑干扰的第 1 层时(即最近的干扰,次近的干扰不考虑),邻近 P 点有两个干扰基站,如图 3.1 所示。假定所有基站位于小区的中央,这两个基站的距离是$(D-R)$和$(D+R)$,相应的信干比为

$$SIR=\frac{R^{-n}}{(D-R)^{-n}+(D+R)^{-n}} \tag{3.2}$$

在 3 小区再用的例子中,$D=6R$;4 小区再用中,$D=8R$。由式(3.2)计算的结果显示在表 3.1 中(注意信干比用 dB 表示)。

表 3.1　信干比计算,一维情况,单行干扰

再用因子	SIR(dB)	
	$n=3$	$n=4$
3 小区	19.6	27
4 小区	23	32

正如所希望的一样,随着距离变化参数 n 的增加,干扰减小而信干比(SIR)增加了。随着再用因子的增加,干扰也减小了。现在考虑次邻近 P 点的两个干扰小区,基站位置仍在小区中央,P 点的移动台离它们的距离分别是$(2D-R)$和$(2D+R)$。增加这两个干扰之后,信干比为

$$SIR=\frac{1}{\left(\frac{D}{R}-1\right)^{-n}+\left(\frac{D}{R}+1\right)^{-n}+\left(\frac{2D}{R}-1\right)^{-n}+\left(\frac{2D}{R}+1\right)^{-n}} \tag{3.3}$$

注意,已经用小区半径 R(在这里为中心到每个边的距离)来归一化这个等式了。如果次邻近干扰被包括进去,表 3.1 会有怎样的变化呢?比如,在 3 小区再用情况下,$D/R=6$,$n=3$,可以得出 $SIR_{dB}=19.2$ dB,而只有在邻近小区干扰情况下,$SIR_{dB}=19.6$ dB。很显然,只要参数 n 足够大(在这里为 3),计算 SIR 时,距离的影响已经变得非常小(只考虑邻近小区,或最多只考虑次邻近小区即可,别的小区可以忽略)。下一节中,我们将讨论更复杂的情况。

3.3　二维小区簇和信干比

一般的二维小区结构模型,采用六边形形状,且所有的小区尺寸都是相同的。注意到,仅仅有 3 种形状能够铺平一个无限的二维平面:六边形、正方形、三角形。如果考虑基站天线均匀辐射所有的方向或者覆盖特定扇区的情况,六边形模型是最有用的。它提供了一个非常好的近似于圆的模型,就像所看到的,简单的几何属性使信干比的计算非常容易。在实际系统中,由于天线的方向性、位置、传播条件、地形结构等的不同,小区的模型还是非常复杂的。考虑了地形条件和传播环境的软件包已经被开发出来了。就像前几章所提到的,这本书的主要内容是提供一个简单、基本的规则,以帮助读者了解无线通信系统和为了分析而开发有用的系统模型。标准的六边形结构符合这些规则,这种方法也将在后续章节中被使用(也偶尔使用一

下圆小区结构以简化分析和计算)。然后,我们可以对比从六边形模型获得的结果与圆小区模型所得结果。

因此,现在主要讨论二维的六边形小区结构,图 3.2 展示了一个例子。就像上面简单说明的一样,几个小区簇成一组才能提供频率再用,有效信道也被分割成几组。7 小区再用的情况如图 3.2 所示,数字 1~7 代表了 7 个不同的信道组,每个信道组使用不同的频谱。具有相同编号的小区可以使用同样的信道组,当然这样会造成对其他同频小区的干扰。由图可知,最近的干扰小区间隔为 D_7。一般情况下,一组使用不同频率的小区簇里面包括 C 个小区,其中 C 为

$$C = i^2 + j^2 + ij \tag{3.4}$$

(i, j)为整数 0,1,2,…,表 3.2 列出了小区簇的各种情况。简略地说明一下式(3.4)的正确性。一般在实践中,小区再用的例子为 $C = 3, 7$ 和 12(如图 3.2 所示),AMPS 系统为 7,GSM 为 3 或者 4。图 3.3(a)为 3 小区频率再用图案;图 3.3(b)为 4 小区频率再用图案。最近的干扰之间的距离分别用 D_3 和 D_4 表示。

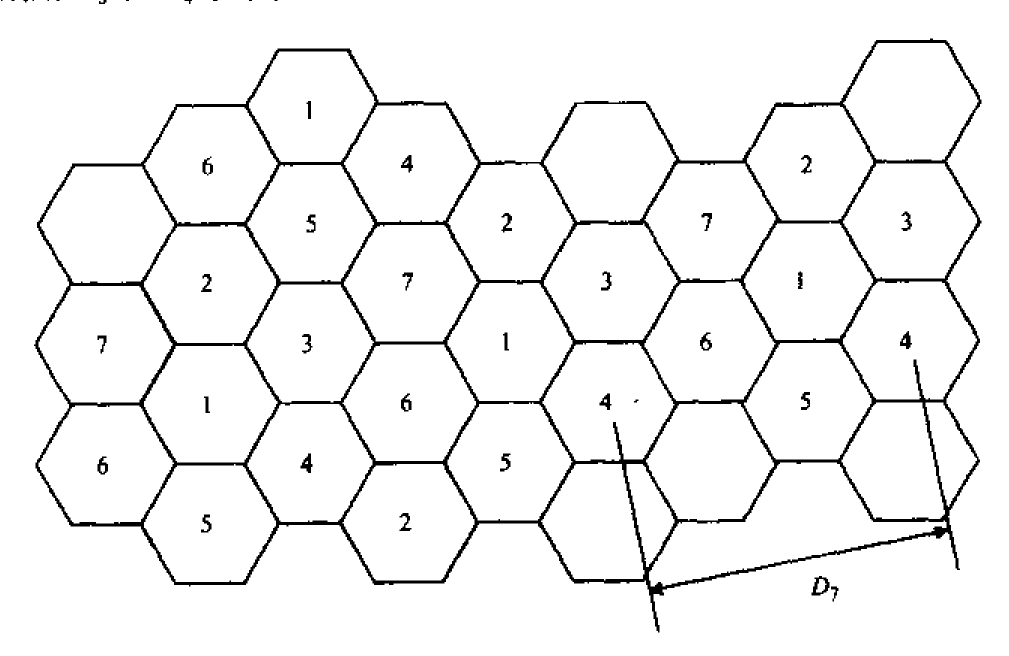


图 3.2 六边形小区结构图,二维

表 3.2 小区簇,六边形小区

C	i	j
1	0	0
3	1	1
4	2	0
7	2	1
9	3	0
12	2	2
13	3	1
16	4	0
19	3	2
21	4	1
27	3	3

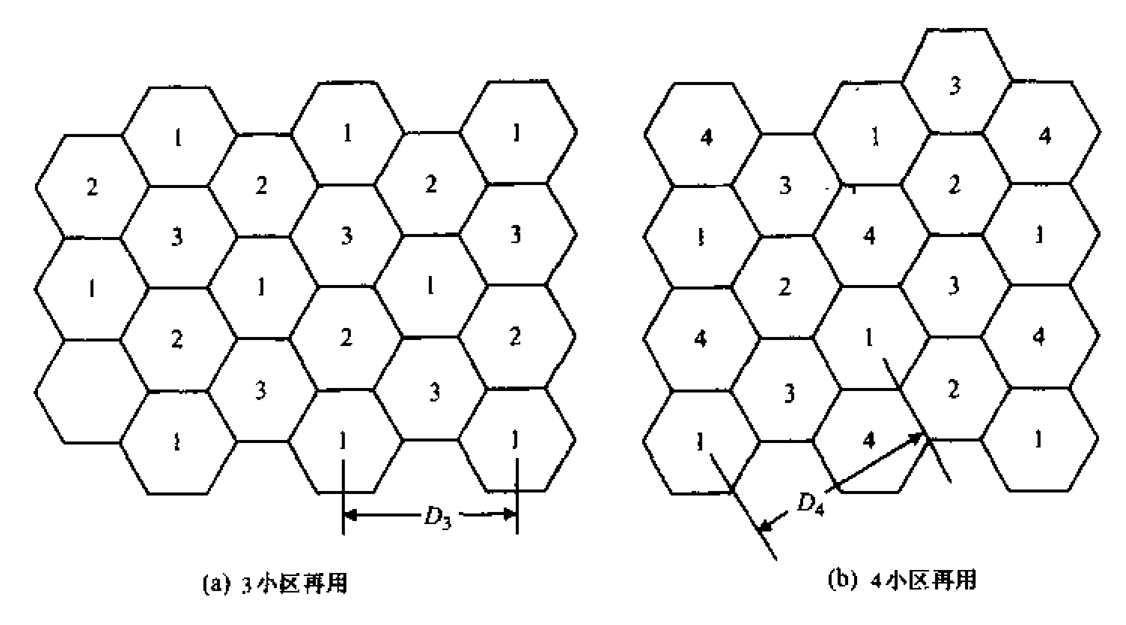


图 3.3　小区再用，六边形小区结构图

考虑一个典型的六边形小区，如图 3.4 所示。从小区中心六边形定点的距离为 R，每个边的长度也是 R。平行边之间的垂直距离为$\sqrt{3}R$，如图所示。小区的面积为 $6(R/2\times\sqrt{3}R/2)=3\sqrt{3}R^2/2=2.6R^2$。注意，和六边形有同样半径 R 的圆的面积为 $\pi R^2=3.14R^2$。六边形内部套一个圆（即圆和六边形的边相切），半径为$\sqrt{3}R/2$，面积为 $3/4\pi R^2=2.36R^2$。这两个圆的面积分别作为六边形面积的上下界。在以后的章节中，将看到有些计算采用圆而不用六边形模型，因为这样可以简化分析。

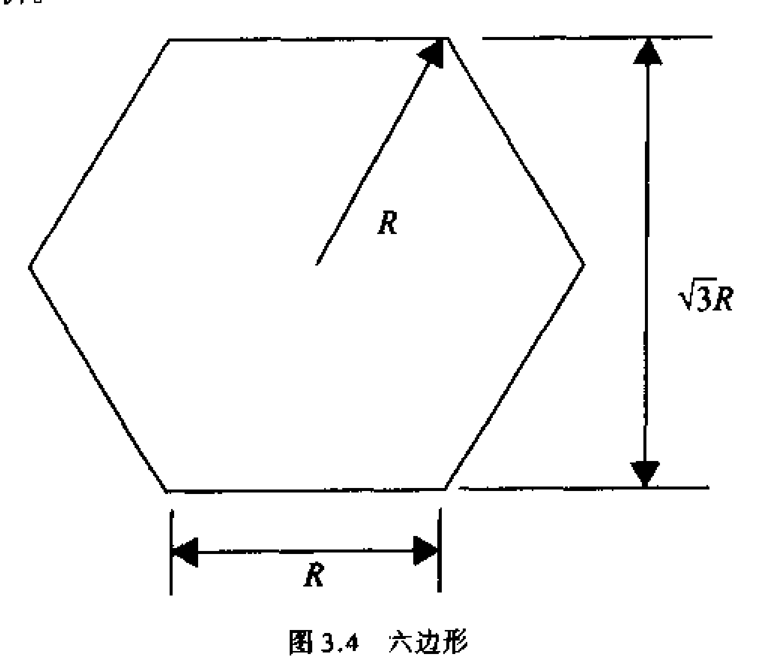


图 3.4　六边形

给定六边形半径 R，可以计算出两个最接近的干扰小区之间中心点的间隔，如图 3.2 和图 3.3 所示，$D_3=3R=\sqrt{3\times3}R$；$D_4=2\sqrt{3}R=3.46R=\sqrt{3\times4}R$；$D_7=4.58R=\sqrt{3\times7}R$。一般情况下，由式(3.4)给出的一组小区数量 C 得到 $D_C/R=\sqrt{3C}$。当一组中小区的数量增加，干扰小区之间的间隔也随之增加，从而减小了干扰。但是，分配给每个小区的信道数量将会减

少，就像以前提到的，降低信道再用的效率也是引入小区结构的原因。再看 AMPS 和 IS-54/136 的例子。由 N 个小区组成了一个地址区域。对于该区域，可使用的频率信道的数量为 $832N/C$，则每个小区可使用信道的数量为 $832/C$。例如，7 小区再用系统为 $832/7\approx118$ 个信道；3 小区再用系统为 $832/3\approx277$ 个信道，这已经在前面一维的例子中提到了。

我们已经验证完了式(3.4)和 $D_C/R=\sqrt{3C}$，首先看后者，使用 Yacoub(1993)[①] 的著作中的方法，从图 3.2 我们可以看到，由 7 个小区组成的簇其本身也近似于一个六边形，如图 3.5所示。这个大六边形的平行边之间的距离为 D_7，如图 3.2 所示。一般情况下，这个值等于相邻干扰小区距离 D_C，如图 3.5 所示。这个六边形的半径 $R_C=D_C/\sqrt{3}$。它的面积是 $A_C=3\sqrt{3}R_C^2/2=\sqrt{3}D_C{}^2/2$。在一个簇内小区的数量由簇面积和小区面积之比得到。小区面积 $a=3\sqrt{3}R^2/2$。已知面积比例，得到 $C=A_C/a=D_C^2/3R^2$，由此得到 $D_C/R=\sqrt{3C}$。

式(3.4)是由六边形平铺的二维空间的数学特性。要解释这个特性，如图 3.6 所示。六边形的中心位置由两个轴 u 和 v 的交点表示。v 轴是垂直方向的，u 轴与水平方向成 30°夹角，起始点位于某个六边形的中心点，标注为(0,0)，轴 u 和轴 v 在此点相交，如图所示。D 是任意其他六边形中心点(u,v)到点(0,0)的距离，由毕达哥拉斯定理得到

$$D^2=(v+u\sin30°)^2+(u\cos30°)^2=u^2+v^2+uv \tag{3.5}$$

依次往上递增，u 轴和 v 轴的增量都为$\sqrt{3}R$ 的倍数，即六边形平行边之间的距离。这样，$u=i\sqrt{3}R$，$v=j\sqrt{3}R$，$i,j=0,1,2,3,\cdots$所以，由式(3.5)可以得到

$$D^2=3R^2(i^2+j^2+ij) \tag{3.5a}$$

这个等式和式 $C=D_C^2/3R^2$ 联立，得到式(3.4)。

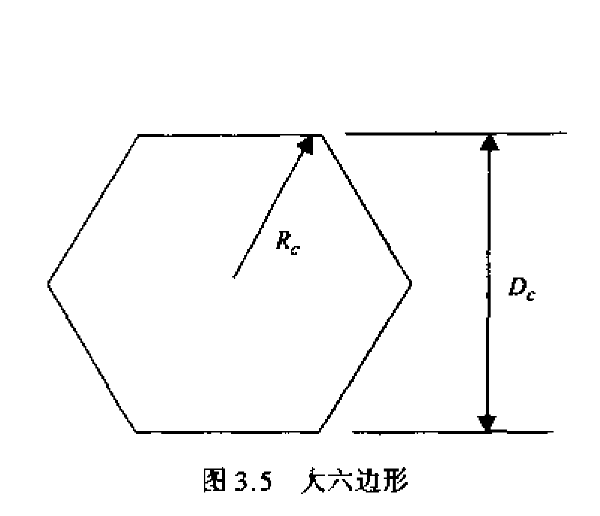


图 3.5 大六边形

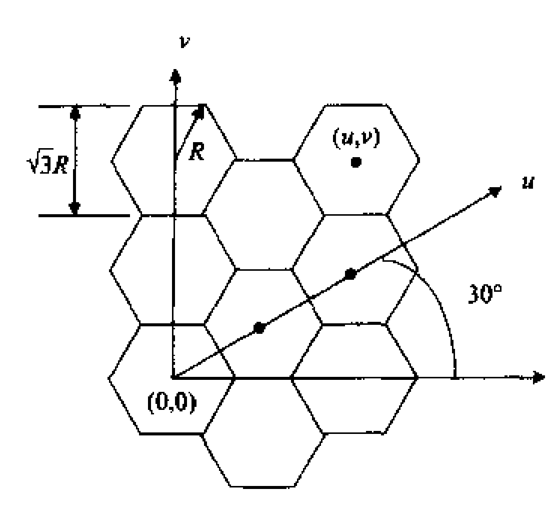


图 3.6 六边形平铺结构

已经介绍了二维小区的情况，六边形平铺结构，第一层小区干扰等，现在来计算二维小区的信干比。从图 3.2 和图 3.3 可以看到，给定一个小区，将有 6 个小区包围着它。这 6 个干扰在图 3.7 中已经标示出来。实际的干扰位置主要依赖于小区簇的大小 C。为了简便，假设沿着六边形间隔为 60°的 6 个轴往下走，可以看出，如果移动台在 P 点，将得到最差的信干比，如图 3.7 所示。此点距其本身基站的距离为 R，距其最近的干扰基站的距离在 1 点，距离大概为

① Yacoub, M. D. 1993. *Foundations of Mobile Radio Engineering*, Boca Raton, FL, CRC Press.

$(D-R)$。距 4 点(也为第 1 层干扰,即邻近干扰 6 个点中的 1 个)的基站距离最远,为$(D+R)$。距其他 4 个基站的距离两两相同(点 2 和点 6 是对称的,点 3 和点 5 是对称的),近似认为它们离移动台的距离都是 D。由 $D_C/R=\sqrt{3C}$, $D/R\geqslant3$, $C\geqslant3$,得到一个近似的信干比为

$$SIR\approx\frac{1}{\left[\left(\frac{D}{R}-1\right)^{-n}+\left(\frac{D}{R}+1\right)^{-n}+4\left(\frac{D}{R}\right)^{-n}\right]} \tag{3.6}$$

使用式(3.6)得到的近似信干比列于表 3.3 中,具体的计算留给读者。

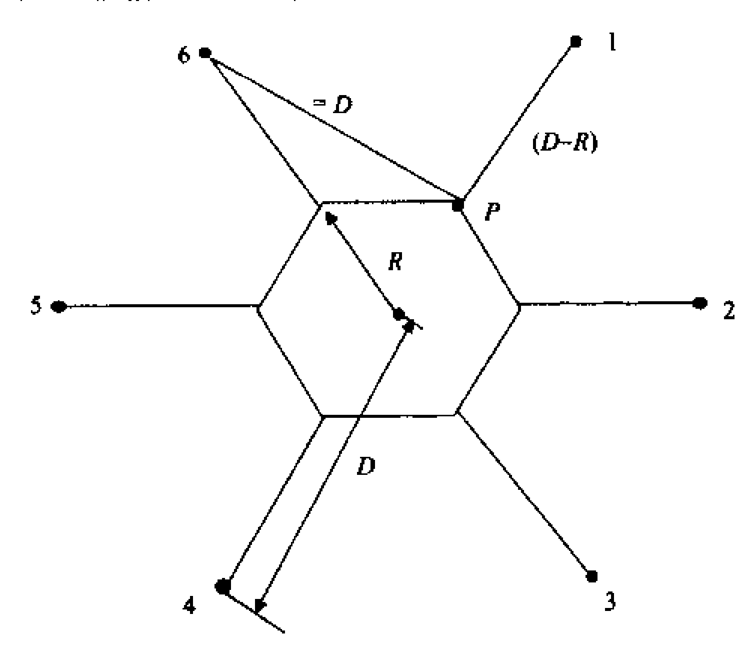


图 3.7　六边形小区的 6 个干扰源

表 3.3　使用式(3.6)得到的近似信干比

C	D/R	SIR		SIR_{dB}	
		$n=3$	$n=4$	$n=3$	$n=4$
3	3	3.4	8.5	5.3	9.3
7	4.58	14.3	62.5	11.6	18
12	6	30.3	167	14.8	22.2

得到式(3.6)和表 3.3 之后,这个近似效果如何呢? G.L.Stüber(1996:468)[①] 已经得到了一个近似值,他使用的干扰距离为$(D-R)$, D 和$(D+R)$,每个点有两个干扰源,其结果为 $C=3$。当 n 分别为 3 和 4 时,信干比为 2.8 和 6.1,相比较表 3.3 中的值为 3.4 和 8.5。$C=7$,当 n 分别为 3 和 4 时,信干比为 13.2(11.2 dB)和 55.6(17.5 dB),而近似值为 14.3(11.6 dB)和 62.5(18 dB),两者仅有0.5 dB的差别。Rappaport(1996:39)[②] 使用了另一个近似值,其中在$(D-R)$的距离处有两个干扰源,在$(D-R/2)$, D, $(D+R/2)$, $(D+R)$处各有一个干扰源。得到的结果更为保守一点,$C=3$, $n=3$时信干比为 2.6, $n=4$ 时为 5.8。但是注意到,当以分贝为单位时,它们的区别并不大。当 C 增加时,增大 D/R,各种近似之间的差别可以忽略掉。

① Stüber, G.L. 1996. *Principles of Mobile Communications*, Boston, MA, Kluwer Academic Publishers.

② Rappaport, T.S. 1996. *Wireless Communications, Principles and Practice*, Upper Saddle River. NJ, Prentice-Hall PTR.

这些结果需要一个 18 dB 的 SIR_{dB}，IS-54/136(数字 AMPS 或 D-AMPS)应该使用 7 小区再用表。实际中，使用 $C=7$，每个小区有 3 根天线，各自覆盖 120°的扇区，由此可以获得 18 dB 的信干比，因为采用了 3 个扇区，小区干扰被减小了 3 倍(干扰为原来的 1/3，每个扇区只出现了两个第 1 层干扰，具体的计算留给读者)。GSM 系统需要 7 dB 的信干比，主要通过引入跳频(发送频率以一个伪随机方式在不同的频率点变化)来减小干扰。由表 3.3，系统可以使用像 3 这样小的小区簇，而传播消耗因子 $n=4$。

3.4 流量控制：爱尔兰厄兰性能和小区尺寸

一旦由系统给定的信干比决定的再用参数 C 和小区簇的大小确定，有效的信道数量也就确定了。例如，第一代 AMPS 系统每个小区提供 $832/C$ 个信道，这已经在前面提到过。当 $C=7$ 时，系统每个小区提供 118 个信道。在以后的章节中我们将看到，第二代数字 AMPS 系统和 IS-54/136 将提供 3 倍的信道。在这些系统里，3 个用户使用同一个频率信道，采用 TDMA 技术的时隙结构，这将在以后叙述。给定每个小区的有效信道数量，能确定小区的地理范围和尺寸。在一个给定区域和系统性能的情况下，主要考虑流量，而不是信干比，信干比主要是针对每个特定的用户而言的。

在这部分，我们将介绍一下体现了系统性能和用户满意度的方法之一——呼叫阻塞概率。在这本书后续章节中，也考虑了系统性能的其他指标，比如切换时的系统掉话率，对于第三代系统中的数据包分组交换技术，要测量丢包率和包延时。历来，呼叫阻塞率一直是电话系统性能指标的主要测量方式。简单地说，它是表征一个用户打算通话但系统告知繁忙的可能性，这点主要依靠系统同时容纳的信道数量和此系统期望的通信流量。在无线通信领域，给定一个小区，呼叫阻塞概率的计算依赖于可以同时处理的有效信道数量以及此小区期望的流量，对此，我们已经讨论过。这里指出的呼叫阻塞概率一般取值在 1% ~ 5%之间，应该在小区期望的流量上加一个限制，这可以通过正确地调整小区尺寸来做到(需要说明一点，在某给定时间段的通信总流量就是在一段时间内的繁忙时间，时间段一般为一周)。在下面，定义了通信总流量，或者流量密度，或者流量负荷。使用了一个简单的模型，用其说明阻塞概率与它的关系。

流量密度被定义为每单位时间内呼叫的平均数量，乘以平均呼叫长度，令 λ 为每单位时间内的平均呼叫数量，$1/\mu$ 是平均呼叫长度。流量密度或者负荷 A 为 λ/μ，用爱尔兰厄兰进行单位表示(A.K.Erlang 是一个伟大的丹麦工程师，早在 20 世纪，他就定量地表示出了流量工程)。举个例子，假如平均每分钟有 50 次呼叫，每次呼叫如果成功的话，持续通话 3 分钟，流量负荷是 150 爱尔兰厄兰。一般的统计模型假设，每次呼叫和呼叫到达服从泊松分布，平均到达速率为 λ。呼叫时间长度在单一模型中被假设为服从指数分布，平均值为 $1/\mu$。一个简单的分析表名，拥有 N 个有效信道，呼叫阻塞概率 P_B，也就是爱尔兰厄兰 B 式为

$$P_B = \frac{A^N/N!}{\sum_{n=0}^{N} A^n/n!} \tag{3.7}$$

(具体的推导见第 9 章。为有线电话而设计的通信流量模型在 Schwartz(1987)[①] 的著作中有比

① Schwartz, M. 1987. *Telecommunication Networks: Protocols, Modeling, and Analysis*, Reading, MA, Addison-Wesley.

较详细的论述,它证实了这个式子独立于呼叫到达——呼叫长度统计特性。该式通过对呼叫到达和呼叫长度统计特性各自使用泊松分布或指数分布进行假设,可以很容易地推导出来。这就是第9章所做的事情)一个 P_B 为纵轴和 A 为横轴的表格或曲线图说明,$P_B = 10\%$ 时,$N \approx A$。当 $P_B = 5\%$ 时,需要 $N > A$ 个信道[爱尔兰厄兰 B 式的详细表格在 Lee(1993)的著作中][①]。举个例子,如上所计算的,如果通信负荷 $A = 150$ 爱尔兰厄兰,$P_B = 5\%$ 时,需要 $N = 153$ 个信道。对于1%的阻塞概率,需要169个信道。信道数量只要增加一点,P_B 的降低也是很明显的。反过来,信道数量是固定的,爱尔兰厄兰负荷中的阻塞概率会怎样改变呢?取 $N = 100$,且不变。若保持 $P_B \leqslant 1\%$,则只允许至多84个爱尔兰厄兰的负荷。如果 A 增加13%,到95爱尔兰厄兰,则阻塞概率将变为5%,增加了5倍。在低的阻塞概率情况下,阻塞概率对于负荷的变化非常敏感。

现在来考虑上面讨论的几个量与给定信道灵活使用之间的关系。如果每个呼叫平均持续了200 s,平均每隔15 min用户进行一次呼叫,有100个有效信道,呼叫阻塞概率为1%,则可以容纳378个用户同时使用(这个数字是怎么得到的?)。如果平均每个呼叫持续了400 s,那么仅仅可以容纳189个用户同时使用。如果每个呼叫平均持续了200 s,但是平均每隔10 min用户进行一次呼叫,那么小区中的用户数量只能有250个。所以,小区电话或终端的使用在决定同时容纳的用户数量时起到了非常重要的作用。

现在来考虑小区尺寸的计算所带来的影响。首先看一个AMPS的例子,回顾前面说到的例子,832个有效信道,7小区再用,每个小区118个信道。要保持阻塞概率 P_B 小于1%,通过使用式(3.7)和基于此等式的表格和图,爱尔兰厄兰负荷为 $A = 101$ 爱尔兰厄兰。例如,每个用户平均每15 min进行一次200 s长的通话,则总的平均呼叫到达速率为 $\lambda = 0.5$ 次/s,在每个小区有 $15 \times 60/2 = 450$ 个用户可以同时通话。现在假设这些用户在小区内为均匀分布,半径为 R 的六边形小区面积为 $3\sqrt{3}R^2/2 = 2.6R^2$,可接受的用户密度为 $173/R^2$ 个/单位面积。例如,首先考虑农村地带,移动终端的密度为每平方千米两个移动台,小区半径 $R = 9.3$ km或5.6 mile。郊区移动终端的密度为100个/km^2 移动台,相应的半径为 $R = 1.3$ km。城市中的密度为1000个/km^2 移动台,相应的半径为 $R = 0.26$ km,也就是800 in。同样的计算方法也可以应用在第二代系统上,如AMPS的数字化系统、D-AMPS系统和IS-54/136系统。前面已经说过,它们的处理能力是原先AMPS系统的3倍,在每个频率信道上,如果使用TDMA技术,则有3个时隙有效。这个系统的小区半径可以设计为AMPS系统的 $\sqrt{3} = 1.7$ 倍。欧洲第二代移动通信系统GSM有124个频率信道,每个200 kHz宽,总带宽25 MHz。这点在以后的章节中会看到,每个频率信道有8个独立的TDMA时隙被分配给用户信道,这样就有992个信道,算上再用参数 C,每个小区有3330个有效信道。如果 $C = 4$,也有248个有效信道。相似的,假设为六边形模型,GSM系统小区半径的计算和AMPS系统、D-AMPS系统的一样。读者可以自己计算一下。

有些时候如果要得到一些更精确的计算结果的话,必须要考虑实际情况,比如移动台在小区内不是均匀分布的。尤其是在一些都市区域,移动台仅仅覆盖一些大街。举个例子,纽约的曼哈顿区就是典型的网格结构(Jakes,1974:Chap.7)。例如,大街间隔为 w,S 为移动台之间的间隔,可以计算在半径为 R 的小区内的移动台的数量。通过一个简单的计算,得到移动台的

---

① Lee, W.C.Y. 1993. *Mobile Communications Design Fundamentals*, New York, Wiley Interscience.

数量 M 为

$$M = \alpha R^2/wS \tag{3.8}$$

对于六边形，$\alpha = 10.4$（考虑一个面积为 A 的小区，分割成一个二维的曼哈顿区样式的网格，大街之间间隔为 w，如图 3.8 所示。大街总长 L 近似为面积为 w^2 的小正方形的数量乘以每个小正方形的周长 $4w$，这样 $L \approx 4wA/w^2 = 4A/w$。对于六边形来说，$A = 2.6R^2$，则 $L \approx 10.4R^2/w$。移动台间隔为 S，发现 $M = L/S = 10.4R^2/wS$。其他的几何形状只在值 α 上有所不同）。

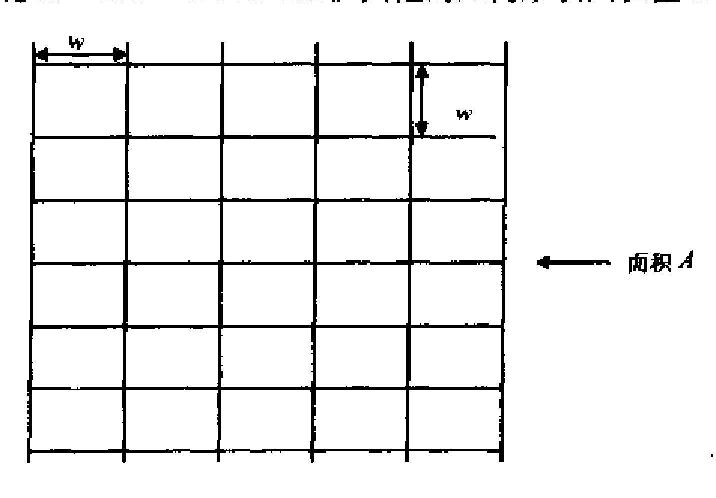


图 3.8 曼哈顿街区网格结构

举个例子，w 为 10 m，当移动台之间的间隔 S 为 30 m 时，每个移动台平均每隔 15 min 呼叫一次，每次 200 s 长，用户呼叫的统计特性和前面使用的例子一样。每个移动用户的爱尔兰厄兰负荷为 $200/15 \times 60 = 0.22$ 爱尔兰厄兰，总负荷为 $0.22M$。考虑前面提到的相同的 AMPS 的例子：再用因子 $C = 7$，每个小区有 118 个有效信道。如果要保持呼叫阻塞概率为 1%，可以容纳总共 101 爱尔兰厄兰的负荷。和 $0.22M$ 相等，在这种环境下，小区的半径为 $R = 400$ m $= 0.4$ km。

3.5 信号的概率计算

之前我们已经仔细考虑了平均功率问题，在 3.1 节和 3.2 节中，计算了平均信干比，使用信干比测量了小区的系统性能。但是，我们忽略了第 2 章中讲过的衰落现象，在小区范围内信号功率与噪声功率都是随时间发生统计性变化的。在一个小区内，更合理的性能测量应该为在这个点所处移动台上行和下行的功率大于干扰功率的概率。可以计算在移动终端实际的信号功率与干扰功率比的概率，测量信干比（注意，这里不是平均信干比）大于某个指定阈值的概率。在一些文献中，已经得到了这些计算结果。我们将在这本书的后面研究有关概率性能测量的问题。这些计算比较复杂，在这些问题上的讨论，主要集中在较为简单的问题上：移动终端接收信号功率超过指定阈值的概率，然后反过来，用这个值来决定小区的哪个部分能接收到可靠的服务。使用第 2 章中阴影衰落模型或对数正态分布衰落模型来得到期望的结果，这些方法都是从 Jakes（1974：125 ~ 128）的著作中引入的。

回顾第 2 章，注意到，经过许多波长的距离到达接收端的大尺度信号衰落可以建模为对数正态分布模型：用 dB 表示的接收信号的值随机地改变，且服从高斯分布。图 2.3 描绘了作为发送接收距离的函数，其典型接收信号功率变化的曲线图。由于空间中有许多的散射点，这些

散射点造成了平均功率的随机变化。由于大尺度衰落,在距离发射端为 d 的接收机处,实际的典型散射平均功率的测量建模为对数正态分布。这里重复式(2.4),接受信号功率为 P_R

$$P_R = \alpha^2 10^{x/10} g(d) P_T G_T G_R \tag{3.9}$$

α^2 是由于多径造成的小尺度瑞利衰落和莱斯衰落的因子,这里忽略多径效应。$g(d)$ 表示距离,形式为 d^{-n},在这一章中主要用于计算信干比。其他的参数是发送端的发射功率和接收端的大线增益。这里集中讨论函数 $10^{x/10}$。

这样,考虑用 dB 表示的接收功率。忽略小尺度衰落模型,设定 $\alpha^2 = 1$, $P_{R,\mathrm{dB}} = x + P_{R,av,\mathrm{dB}}$,其中,$P_{R,av,\mathrm{dB}} = 10\lg(g(d)P_T G_T G_R)$。简化这个表达式,令 $P \equiv P_{R,\mathrm{dB}}$, $P_{av} \equiv P_{R,av,\mathrm{dB}}$,则用 dB 表示的随机变化的接收信号功率 p,其平均值为 p_{av},是服从高斯分布的,也用 dB 表示。重复式(2.9),它的概率密度分布函数为

$$f(p) = \frac{e^{-(p-p_{av})^2/2\sigma^2}}{\sqrt{2\sigma^2}} \tag{3.10}$$

标准方差 σ 取值在 6~10 dB 之间,这点在第 2 章中说明过。以分贝表示的接收信号功率的高斯分布模型是如何使用的呢?例如,在移动台一方,可接受的接收信号功率的阈值为 p_0 dB,给定 p_{av} 和 σ,都用 dB 表示,能计算接收信号功率大于某个数值的概率,这个概率 $P[p \geqslant p_0]$ 由下式给出

$$\begin{aligned} P[p \geqslant p_0] &= \int_{p_0}^{\infty} \frac{e^{-(p-p_{av})^2}}{\sqrt{2\pi\sigma^2}} \mathrm{d}p \\ &= \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{p_0 - p_{av}}{\sqrt{2}\sigma}\right) \end{aligned} \tag{3.11}$$

其中,差错函数 $\mathrm{erf}(x)$ 定义为

$$\mathrm{erf}(x) \equiv \frac{2}{\sqrt{\pi}} \int_0^x e^{-x^2} \mathrm{d}x \tag{3.12}$$

在许多书中,差错函数依据概率和统计特性已经被列成了表格,在许多软件中也包含这个函数的专门的软件包。图 3.9 和式(3.11)都在说明差错函数。注意,给定平均接收功率 p_{av} 和标准方差 σ,式(3.11)也表示超过指定接收功率 p_0 的概率。

举个例子,令 $p_0 = p_{av}$,接收信号大于阈值的概率为 50%,如阈值 p_0 大于 p_{av} 10 dB。平均接收功率是 1 mW,阈值设定为 10 mW,阴影衰落特性的标准方差也是 10 dB,超过阈值的概率为 $\frac{1}{2} - \frac{1}{2}\mathrm{erf}(1/\sqrt{2}) = 0.16$。给定 $\sigma = 10$ dB,接收功率大于平均值 p_{av} 低于 10 dB 的概率为 0.84。

现在考虑建模阴影衰落效应为对数正态分布的另外一个应用(Jakes,1974)。如上所述,在一个小区内,可接受的接收功率在距离基站 R 的地方,被设定为平均功率值 p_0。现在,问题是小区中哪一部分服务区域是可靠的,它的接收功率大于或等于 p_0? 注意到由于 $(r/R)^{-n}$, $(r < R)$ 使得功率增大,如果没有阴影衰落的影响,小区半径内所有的区域都可以接收到满意的服务。再看以基站为圆心的半径为 R 的圆,为了简便起见,假设使用全向天线(显然,方向性天线能够得到更高的增益,用以改变此位置的接收质量)。我们称部分有用的服务区域为 F_u,所以

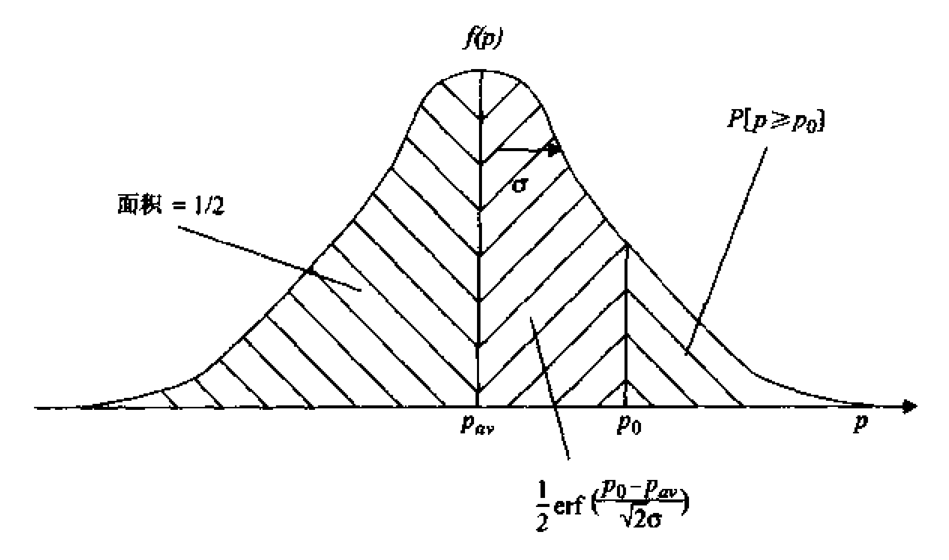


图 3.9 接收信号(dB)大于某个阈值的概率

$$F_u = \frac{1}{\pi R^2} \int_{over \pi R^2} P(p \geq p_0) \mathrm{d}A \tag{3.13}$$

其中，$\mathrm{d}A = 2\pi r \mathrm{d}r$。在图 3.10 中，看到了与式(3.13)计算相关的几何图形。既然已经在 $r = R$ 的地方取 $p_0 = p_{av}$，则平均接收功率依据距离 r^{-n} 变化，以分贝表示的平均接收功率 p_{av} 可以被写成 $p_{av} = p_0 - 10n\lg(r/R)$。将此式插入到式(3.11)中可得

$$P[p \geq p_0] = \frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{10\lg r/R}{\sqrt{2}\sigma}\right) \tag{3.14}$$

(很明显，可以近似地假设 r^{-n} 中的参数 r 包括 $r = 0$)式(3.13)被改为

$$F_u = \frac{1}{\pi R^2} \int_0^R 2\pi R \left[\frac{1}{2} - \frac{1}{2}\mathrm{erf}\left(\frac{10n\lg r/R}{\sqrt{2}\sigma}\right)\right] \mathrm{d}r \tag{3.15}$$

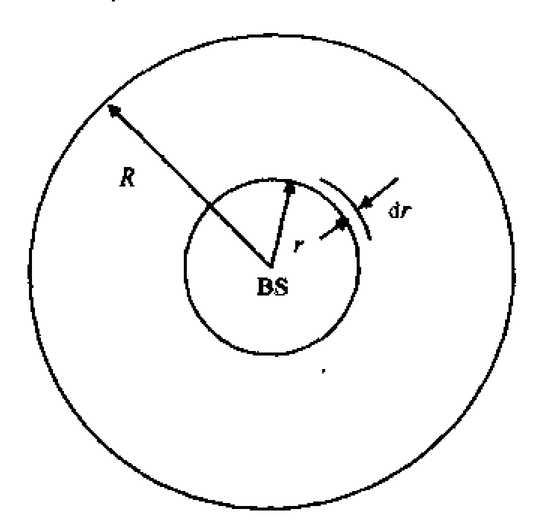


图 3.10 有用服务区域的计算

通过转化 lg 为 ln，简化式(3.15)，定义一个参数 $b \equiv 10n\lg(e/\sqrt{2}\sigma)$，通过式 $e^x \equiv (r/R)^b$ 或 $x = b\ln(r/R)$，定义一个变量 x，用 x 取代积分变量 r，得到

$$F_u = \frac{1}{2} + \frac{1}{b}\int_0^{\infty} e^{-\frac{2x}{b}} \operatorname{erf}(x)\,dx \tag{3.15a}$$

通过用式(3.12)中定义的积分取代 erf(x),可以直接对式(3.15a)进行积分。合成的双重积分可以通过使用交换积分顺序等方法来得到。通过上面 b 的实际数值,在取代了参数 b 以后,得到如下结果:

$$F_u = \frac{1}{2} + \frac{1}{2}e^{\left(\frac{\sigma}{3n}\right)^2}\left[1 - \operatorname{erf}\left(\frac{\sigma}{3n}\right)\right] \tag{3.15b}$$

[看到结果仅仅与 $\sigma/3n$ 有关,还可以将 $10\lg(e/\sqrt{2}) = 3.13$ 近似为 3]。

举个关于此表达式使用的例子,比如 $n = 3$(即由距离导致的 r^{-3} 的功率变化)和 $\sigma = 9$ dB,得到 $F_u = 0.71$,这表明在以此基站为圆心,半径为 R 的小区内,有 71% 的地方可以得到大于阈值(此阈值设定为距基站 R 远的平均接收功率)的信号,而沿着半径为 R 的圆周,有一半地方可以接收到阈值以上的信号。如果 $n = 4, \sigma = 6$ dB 则 $\sigma/3n = 0.5, F_u = 0.83$(Jakes, 1974; Fig.2.5-1, p.127)。由基于 σ/n 的式(3.15b)可见,n 增加(σ 减少),F_u 增加。我们所希望看到的变化是,当一个人朝着基站方向移动时,随着 n 的增加,接收功率随之增加得更快,同时收到阈值以上信号的区域也增加。同样,减少 σ,依据测量阴影衰落变化导致的平均接收功率的变化,有效区域覆盖也会增加。

更通俗地说,在距离基站 R 的地方,令可接受的接收信号功率 p_0 和平均功率 $p_{av}(R)$ 不同,定义一个参数 $a \equiv [p_0 - p_{av}(R)]/\sqrt{2}\sigma$,然后,我们(Jakes, 1974)可得到 F_u 为

$$F_u = \frac{1}{2}\left\{1 - \operatorname{erf}(a) + \exp\left(\frac{1-2ab}{b^2}\right)\left[1 - \operatorname{erf}\left(\frac{1-ab}{b}\right)\right]\right\} \tag{3.16}$$

这里,$b = 3n/\sigma$ 和以前定义的参数一样。

当阈值 p_0 下降到平均值 $p_{av}(R)$ 以下时,也就是说,参数 a 为负数,F_u 增加。概率 $P[p_{av}(R) > p_0] = 0.7$ 增加,小区内将会有更多的地方得到比阈值大的信号。有效的服务面积增大了。举个例子,比如阈值 p_0 减小,以便 $P[p_{av}(R) > p_0] = 0.7$,而不是前面的 0.5。再次令 $n = 3, \sigma = 9$ dB,则在上面的例子中,$b = 1$。此时,(Jakes, 1974; Fig.2.5-1)有效小区从先前的 0.71 可增加到 0.84。

习题

3.1 (a) 验证表 3.1 中的数据。

(b) 对比包括第 2 层干扰在内的信干比与只有第 1 层干扰的信干比的区别。证明:当第 2 层干扰在内,$D/R = 6, n = 3$ 时,$SIR_{dB} = 19.6$ dB。

3.2 (a) 考虑一个半径为 R 的六边形小区,如图 3.4 所示。证明:小区的面积是 $3\sqrt{3}R^2/2 = 2.6R^2$。说明由于小区再用参数 C 的影响,两个干扰距离最近的小区中心间隔为 $D_c/R = \sqrt{3C}$。

(b) 参考图 3.6,填写由式(3.4)得到的结果的表格。

3.3 (a) 验证表3.3中的数据。

(b) 观察7小区再用中的表3.3中的数据，将这些数字与课本中建议的两个其他模型近似的结果进行对比，你能提供一种新的方法来近似干扰的位置吗？

3.4 一个移动台，距离小区中心 $D=3.46R$，假设此移动台受到邻近的移动台的联合信道干扰，请计算最差的上行信干比。使用最简单的路径损失模型 $g(d)=1/d^4$。当 $n=3$ 时，重复上述步骤。

3.5 计算下面系统的爱尔兰厄兰负荷

1. 呼叫平均持续200 s，每分钟平均100次呼叫。
2. 小区内有400个移动用户，每个用户平均每隔15 min呼叫一次。呼叫平均持续3 min。
3. 在2中，用户的数量增加到500，每个用户平均每隔20 min呼叫一次，呼叫平均持续6 min，计算结果。

3.6 考虑一个支持100个信道的移动通信系统，呼叫阻塞概率为1%，用户平均每隔10 min呼叫一次，呼叫平均持续10 min。例如，系统在都市中，每平方千米500个用户，如果使用六边形拓扑结构，请计算所需的小区半径。如果呼叫平均持续4 min，小区尺寸有什么变化？多长时间呼叫一次？

3.7 路径损失建模为

$$
\begin{aligned}
g(d) &= d^{-1.3} & 0 \leqslant d \leqslant 150\,\mathrm{m} \\
&= 150^{-1.3}(d/150)^{-3.5} & 150\,\mathrm{m} \leqslant d
\end{aligned}
$$

[参见第2章中的式(2.6a)]。

当再用参数 $C=3$ 时，计算最差情况的上行信干比，分两种情况：1、$R=1000$ m；2、$R=100$ m。

3.8 参考本章3.4节，IS-136系统和D-AMPS系统能处理3倍于AMPS系统的通信流量，小区半径也可扩大1.7倍。解释一下，这里面包含了哪些隐含条件。

3.9 考虑一个小区再用参数 $C=3$ 的GSM系统，其中每个小区有330个有效信道。假设允许的爱尔兰厄兰流量负荷有同样的数字（这里鼓励读者自己计算一下由爱尔兰厄兰B式得到的图表，然后得到呼叫阻塞概率）。假设使用六边形模型，在郊区移动台密度为200台/km²，计算小区半径。当 $C=4$ 时，重复上述计算。这些结果如何改变以应用于移动台密度为1000台/km²的都市区域？所有情况都设定呼叫平均持续200 s，每个用户平均每隔15 min呼叫一次。

3.10 参考由式(3.12)推导出的错误函数的表格。假设阴影衰落为对数正态分布，方差为 σ^2，功率变化参数为 n，式(3.15b)代表了可接收到有效服务的小区百分比。当 $n=3$，$\sigma=9$ dB时，这个数字是0.71，当 $n=4$，$\sigma=6$ dB时，此数字增加到0.83。找到这两个数字的其他结合方式，并计算结果。

第4章　动态信道分配和功率控制

第3章中，讨论了复用受限的蜂窝环境下的信道分配，以保持干扰低于一定可容忍的强度。复用受限减少了每个小区可分配的信道数目，也使得利用蜂窝结构带来的系统容量(system traffic capacity)的改善有所减小。人们提出并采用了很多策略以进一步提升系统性能。一种方法就是利用有向天线减少干扰信号的数目，使得期望接收的信号符合要求(Stüber, 2001)[①]。最近的蜂窝系统实际上已经采用了这种方法。另外一种方法就是在给定的复用特性的情况下，减小小区面积，使得在一定几何区域内的小区数量增加，从而带来容量的提高。蜂窝系统呈现出按小区面积分层的结构，如宏小区、微小区、微微小区。宏小区面积较大；在拥挤的城市环境中负荷量较大，可采用微小区以减小小区面积(基站天线和附属的发送接收系统可以安装在较低的高度，如街道的灯柱上)；微微小区主要在室内蜂窝系统中使用。微小区的一个主要问题是切换较频繁，需要更有效的处理切换。这个问题将在第9章中进行讨论。在宏小区基础上覆盖微小区的结构也已经被提出，以用来提高业务处理能力(Stüber, 2001)。

还有一些改善系统性能的方法，如动态信道分配DCA，它能降低呼叫阻塞率，提高业务吞吐量，如降低干扰的功率控制。我们将在本章讨论这两种技术。一些仿真结果表明，相对于单一地使用其中一种技术，动态信道分配和功率控制的结合能够极大地改善系统性能(Lozano和Cox, 1999)[②]。

在4.1节中我们将介绍动态信道分配，主要针对一种分配策略以简化讨论。在一些所谓低阶(low-tier)的无线系统中，采用的是一些简单的动态信道分配策略。这些系统，如欧洲的数字无绳电话系统(DECT)和日本的个人手持电话系统(PHS)，与本书后面章节将讨论的高阶(high-tier)第二代移动通信系统不同(Cox, 1995[③]; Goodman, 1997[④]; Rappaport, 1996[⑤])，它们都是主要针对步行通信业务设计的低功耗的系统。第10章中讨论的第三代移动通信系统和未来的移动，预计将会在更大程度上利用动态信道分配技术。无线移动系统中的功率控制将在4.2节中讨论。列举了一些功率控制的方法以引出这个课题，并提供了相关的控制过程的实例。在第3章中，我们已经证明了控制发送功率使得保持一定的信干比的必要性。在后面的章节中，对码分多址系统的讨论将会发现功率控制是必不可少的要求。

4.1　动态信道分配

在前面章节中提到的小区信道分配方法都假定采用了固定信道分配FCA，信道以组为单

---

① Stüber, G. L. 2001. *Principles of Mobile Communications*, 2nd edn, Boston, MA, Kluwer Academic.

② Lozano, A. and D. C. Cox. 1999. "Integrated dynamic channel assignment and power control in TDMA wireless communication systems," *IEEE Journal on Selected Areas in Communications*, 17, 11 (November), 2031 – 2040.

③ Cox, D. C. 1995. "Wireless personal communications: what is it?," *IEEE Personal Communications*, 2, 2, 20 – 35.

④ Goodman, D. J. 1997. *Wireless Personal Communication Systems*, Reading, MA, Addison-Wesley.

⑤ Rappaport, T. S. 1996. *Wireless Communications, Principles and Practice*, Upper Saddle River, NJ, Prentice-Hall PTR.

位,按规定的复用图案固定不变地分配给确定的小区(前面的章节提到的复用策略,是将每 C 个小区组成一簇,把所有可用的信道集合划分为 C 组,每组再分配给每簇中一个固定的小区)。所谓动态信道分配,指的是各种对各小区内用户进行信道分配,使其能够满足动态变化的业务需求的方法。这些方法种类繁多,比如不进行固定的信道分配,将所有的信道集中起来,当需要分配的时候再将信道分配给小区;又如将信道名义上分配给各小区,当需要的时候从临近小区中借取当前不用的信道。但是,这种借用策略需要考虑复用限制。当复用区域发生变化时,借用策略可能变得复杂。采用动态信道分配的策略提高系统容量的可能性在早期的蜂窝系统研究中就已经为人们所认识(Cox 和 Reudink,1972[①],1973[②])。在接下来的若干年中又提出了很多动态信道分配的策略。不管是用途简单的,还是能改善当流量超载时系统的性能的,每种策略都有其各自的特点。超载指的是超过系统设计的额定负荷时的业务强度。引起的原因包括用户数目的增多或者使用时间的延长(用户频繁通话或者通话时间过长)。在前面的章节中,将额定业务负荷定义为在小区信道数目固定的情况下,能够满足理想的小区阻塞率的业务负荷。当负荷超过这个值时,系统负荷越重,呼叫阻塞率越大。当在某个固定区间内时,一种策略的阻塞概率曲线在另一种曲线的下方,则认为这种分配策略较优。

有关动态信道分配策略的概述,连同不同策略间定性的比较,可参见 1996 年 Katzela 和 Naghshineh[③]的论文。如前面所述,动态信道分配在信道分配流程中引入了灵活性。不同的动态信道分配策略在信道分配规则上有所区别。值得注意的是,所有的动态信道分配方案的性能在低、中业务流量条件下优于固定信道分配策略。然而在高负荷情况下,固定信道分配性能较优,原因在于高负荷情况下,几乎没有空闲信道来分配给新的呼叫。此外,下文将要解释的所谓信道锁定,也会使情况恶化,造成性能的损失。方案提出采用混合策略,即在高负荷条件下变为固定信道分配。

在动态信道分配的实现中,当需要时,小区可以从临近小区借取信道。但如果借取了信道,由于复用限制,其他小区就会被锁定而不能使用这些信道。因此,这种动态信道分配中借取的方法会导致信道锁定。在适度的业务负荷下,这不会带来问题,动态信道分配能够调节不同小区间业务的波动。但在重度负荷情况下,信道锁定会引起小区间更为频繁的借取,最终非但不能改善性能,反而会使性能降低。信道锁定可参见图 4.1。如图所示,将信道每 7 个小区为一簇进行划分。信道不能为临近的小区所复用,但可以被更远的小区使用。比如在一种动态信道分配方法中,若 7 号小区中的信道 x 被 1 号小区借走,则在 1 号小区周围的所有小区,连同 7′号和 7″号小区,其信道 x 都被锁定。信道 x 不能被 7′号和 7″号小区周围的小区借取,这就是由于复用限制造成的信道锁定问题,正如上面提到的,在高负荷情况下会削弱动态信道分配的性能。

值得注意的是,这种简单的借取策略过于严格,可以采取"定向锁定"的方法来缓解部分问题。在同一个例子中,仍参考图 4.1,在 7′号小区远离 1 号小区的一侧的所有小区,如 2′号和 1′

① Cox, D. C. and D. O. Reudink. 1972. "Dynamic channel assignment in two-dimensional large-scale mobile communication systems," *Bell System Technical Journal*, 51, 1611 – 1672.

② Cox, D. C. and D. O. Reudink. 1973. "Increasing channel occupancy in large-scale mobile radio systems: dynamic channel assignment," *IEEE Transactions on Vehicular Technology*, VT-22, 218 – 222.

③ Katzela, I. and M. Naghshineh. 1996. "Channel assignment schemes: a comprehensive survey," *IEEE Personal Communications*, 3, 3 (June), 10 – 31.

号，可以在不违反复用限制的情况下借取信道 x。这种定向锁定的方法确实能够改善动态信道分配的性能。

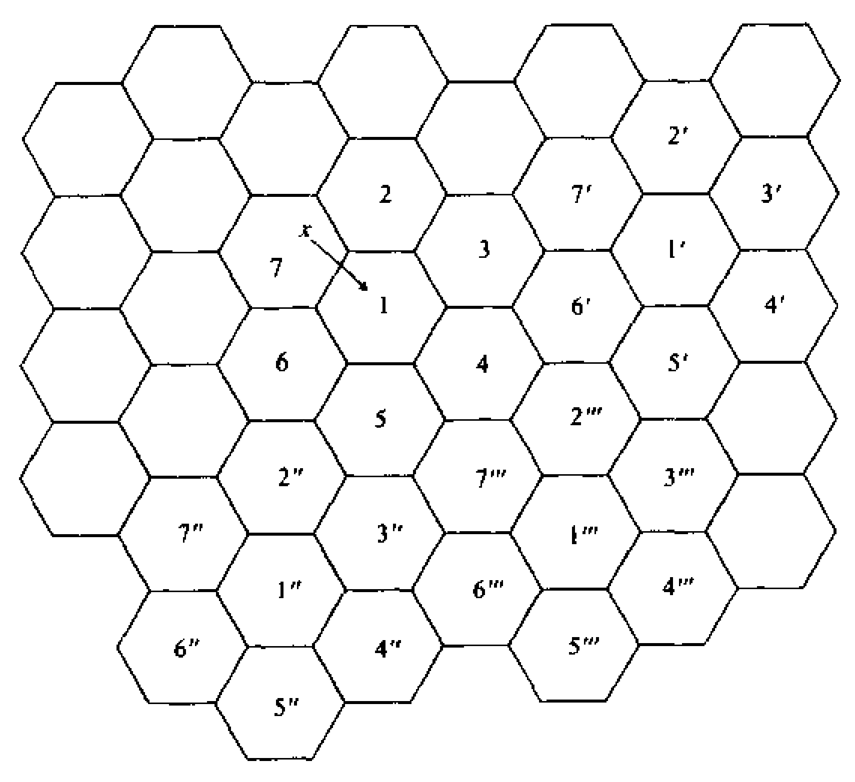


图 4.1 信道锁定

如前面所述，多种信道分配的技术已经在相关文献中被提出(Katzela 和 Naghshineh，1996)。这些方法包括如何选取合适的小区进行信道借取；不同的信道分配规则；将部分信道分配给各个小区，将剩余部分集中起来根据需要自由分配；在小区的外部和内部采用不同的借取策略，等等。本书中不对每种方法逐个介绍，而是集中介绍一种动态信道分配策略，它已被理论和仿真所证明，能够提供比该文献中大多数提到的方法都好的性能。有兴趣的读者可以仔细阅读该著作中提到的其他策略。下面介绍的方法称为定向信道锁定 BDCL 的信道借用的动态信道分配策略，由 Zhang 和 Yum(1989)①提出。这种方法能够进行理论分析，而不仅仅是仿真验证，也使对其作用机制的解释变得简单。BDCL 是定向信道锁定的信道借用(borrowing with directional channel locking)的首字母缩拼词，这种方法利用定向锁定能够获得性能的改善。整个方案主要由 3 个特征定义：锁定方向的确定、信道排序和信道立即重分配。在信道排序中，每个小区分配的信道被标号。标号最小的信道具有分配给下一个本地呼叫的最高优先分配权，标号最大的信道则将最先借给临近小区。本地呼叫将从最低标号开始依次分配；借取信道则从最大标号开始依次减小。信道立即重分配是一种将本小区中通话信道和借取信道紧密排列的策略，包括 4 个部分：① 本地通话排序整理——当其中一个信道的通话结束，而在较高序号的信道中有通话进行，则将后者移至刚释放的序号较低的信道。② 当一个利用本地信道的通话结束，同时有一个利用借取信道的通话正在进行，后者将被移至本地信道，借用的信道将

① Zhang, M. and T.-S. P. Yum. 1989. "Comparisons of channel assignment strategires in cellular telephone systems," *IEEE Transactions on Vehicular Technology*, 38, 4(November), 211－218.

被释放。③ 借用信道排序整理——当其中一个借用信道的通话结束并释放一个借用信道时，若此时在较低序号的借用信道中有通话进行，则将后者移至刚释放的序号较高的借用信道。④ 若在一个干扰小区中某个通话结束信道锁定解除，任何在借取信道或较高序号的信道中进行的通话将切换至该信道。性能研究表明这种方案在适度的负载下性能优于固定信道分配，也如期望的那样，比不使用定向锁定的方案性能要有所改善（Zhang 和 Yum，1989）。

下文中，主要讨论这种方案的一维形式，同时展示如何进行近似的性能分析。将采用 Yeung 和 Yum（1995）[①]的方法。比如，系统中总共有 $2m$ 个信道，每次给交替的小区分配 m 个信道。把第 1 个 m 个信道组成的集合称为集合 a，第 2 个集合称为集合 b。图 4.2 表示的就是整个一维集合中小区 $i-1$，i 和 $i+1$ 的分配情况。假设小区 i 中所有的信道 m 都已被使用。在固定信道分配的情况下，如果新呼叫到达，则它就会被阻塞。阻塞概率由第 3 章中式（3.7）给出，式中 $N=m$。在动态信道分配策略下，如果有可用信道，则 i 能从 $i-1$ 或 $i+1$ 借取。如果它从 $i-1$ 或 $i+1$ 借取信道，对应的 $i+1$ 或 $i-1$ 的信道则被锁定。在采用定向锁定的动态信道分配策略中，如果在这 3 个小区组成的系统中，所有 $2m$ 个信道都在使用，则小区 i 会阻塞新呼叫。因为小区 i 不能从相邻小区借得信道，即使相邻小区有可用信道，也会由于该信道在另一个相邻小区可能会被使用而导致不可用。这方面的解释在下文会有详细的阐述。

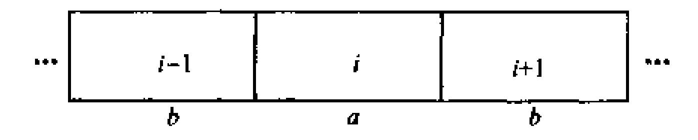


图 4.2 一维蜂窝系统中的信道分配

Yeung 和 Yum 观察到图 4.2 中的 3 个小区（一维系统中任意相邻的 3 个小区）能够从系统中解耦并分别处理，这样就使得这种方案的近似分析和与固定信道分配的比较成为可能。随后，对 3 小区系统的分析给出了阻塞概率性能的上（紧）界。

分析如下（Yeung 和 Yum，1995）：如果标记为 i 的小区已经借出了一个或多个信道给其邻居，则容易证明它不会或者不能从相邻小区借取信道。特别地，若小区 i 已经借了 k 个信道给小区 $i+1$，标号从 $m-k+1$ 到 m（注意，这个小区分配的信道集合 a 其序号是从 1 到 m）。这暗示小区 $i+1$ 有 m 个呼叫正在进行，也就是说，它全部的可分配的 m 个信道都被占用了。此时，小区 $i+1$ 有 $m+k$ 个呼叫正在进行。如果此时小区 i 有一个呼叫到达，则会有两种可能性：① 如果小区 i 中当前呼叫数目小于 $m-k$ 个，那么小区 i 可以直接处理该呼叫；② 小区 i 中当前已经有 $m-k$ 个呼叫数目。在这种情况下，小区 i 和 $i+1$ 中总共使用了 $2m$ 个信道，这样，小区不能再从小区 $i-1$ 借取信道，新的呼叫将被阻塞。根据这个分析结果，信道的借取不会传播。这样，阻塞概率的计算可以通过该组从一维网络中解耦再分别处理来实现。Yeung 和 Yum 证明了按照 3 个小区分组并分别处理进行分析得到的阻塞概率实际上是阻塞概率的一个上紧界。按照他们的方法并且计算一些简单的例子，可确定阻塞概率和负载的函数关系。这些结果会与等同的固定信道分配方案进行性能比较，正如前面提到的，比较表明，动态信道分配在中低负荷情况下性能比固定信道分配好，但在高负荷情况下性能则差一些。

---

① Yeung, K. L. and T.-S. P. Yum. 1995. "Cell group decoupling analysis of a dynamic allocation strategy in linear microcell radio systems," *IEEE Transactions on Communications*, 43, 2/3/4 (February – April), 1289 – 1292.

阻塞概率的分析如下：着重研究图 4.2 中 3 个小区组成的小组。再次假定一维系统中一共有 $2m$ 个信道。其中 m 个分配给小区 i，其他 m 个则分配给小区 $i-1$ 和 $i+1$。由于借取，小区 i 最多可以使用 $2m$ 个信道。如果所有 $2m$ 个信道都分配给了它，那么小区 $i-1$ 和 $i+1$ 由于锁定而没有可用的信道。在其他极端的情况下，小区 $i-1$ 和 $i+1$ 通过借取，各自同时都有 $2m$ 个信道在使用，此时的复用是合乎规则的。而此时，小区 i 没有信道可以使用[即使针对 3 小区的情况分析，两个小区 $i-1$ 和 $i+1$ 各自都能从其左侧和右侧的邻居借取]。这样当小区 $i-1$ 和 $i+1$ 都有 $2m$ 个呼叫进行时，小区 i 没有可用的信道。这些信道直接完全等价。很明显，这 3 个小区构成的小组占用的信道数目不可能超过 $2m$ 的最大值。当且仅当所有 $2m$ 个信道都被该组占据时，到达小区 i 的新呼叫将被阻塞。假定 x_1 为小区 i 中占用的信道数（正在进行的呼叫），x_2 为小区 $i+1$ 中占用的信道数，x_3 为小区 $i-1$ 中占用的信道数。把这些数作为每个小区的状态。三元组(x_1,x_2,x_3)代表 3 小区组的向量状态（占用的信道数目）。可能的状态集合如图 4.3 的三维图形所示。图中考虑了结合定向锁定的信道借取的各种可能性。特别值得注意的是，与平面 $x_1=x$ 相交的图形是由 4 点$(x,0,2m-x)$，$(x,0,0)$，$(x, 2m-x,0)$和$(x,2m-x,2m-x)$构成的长方形。这些点如图 4.3 所示。

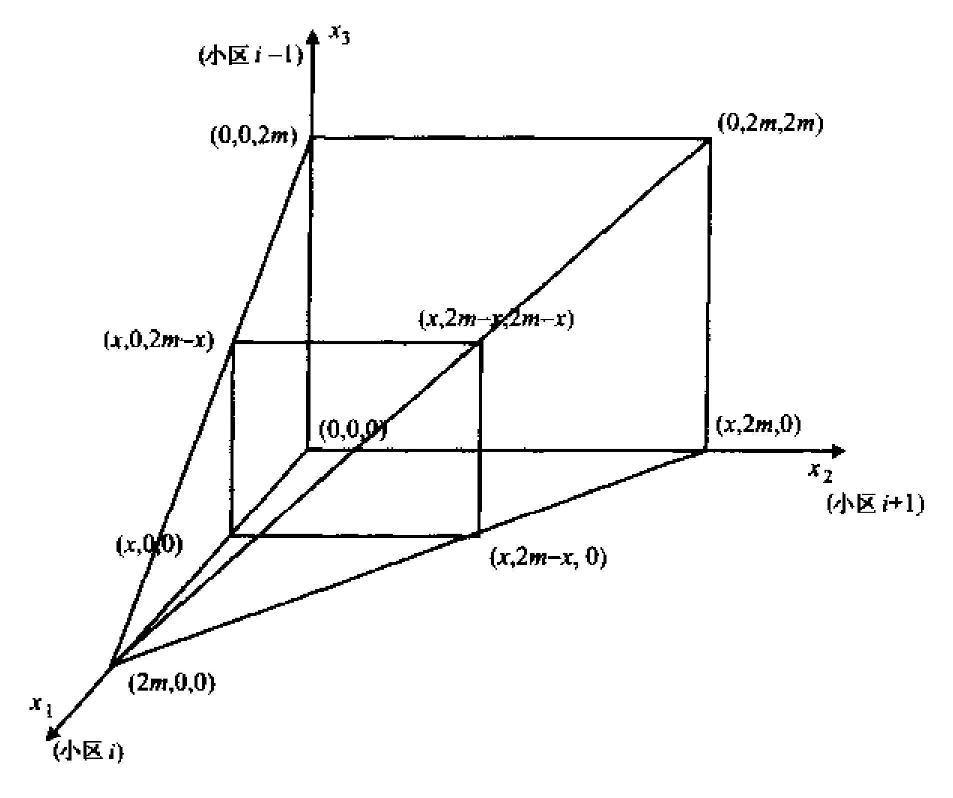


图 4.3　一维定向信道锁定的信道借用策略的可能状态（引自 Yeung 和 Yum，1995）

总共有多少种可能的状态呢？让 x_1 从 $2m$ 变到 0，$2m$ 对应图 4.3 中所显示的只有一种状态的最大点，0 则为 x_1 的最小值。在 $x_1=2m-1$ 时，有 $2^2=4$ 种状态；在 $x_1=2m-2$ 时，有 $3^2=9$ 种状态；直到 $x_1=0$，则此时有$(2m+1)^2$ 种状态，因此，总的可能的状态数为

$$\sum_{k=1}^{2m+1} k^2=(2m+1)(2m+2)[2(2m+1)+1]/6$$

例如，当 $m=2$ 时，总的可用信道数为 $2m=4$，总的可能的状态数目为 55；当 $m=1$ 时，系统内

可用的信道数为$2m=2$，总的状态数为14。在这些可能的状态数中，有一些是阻塞状态，也就是那些在某特定小区中，无论是小区自身还是相邻小区都没有可用信道来接纳多余呼叫请求的状态。例如，一个呼叫请求到达小区，当小区 i 有一个或者多个信道闲置时，呼叫才能被接纳。如果小区 i 没有可用信道，可以从小区 $i-1$ 或 $i+1$ 借取。如果相邻小区也没有信道可用，则呼叫被阻塞。正如前面所提到的，在3小区组中，如果所有 $2m$ 个信道都被使用，则阻塞就会发生。

考虑一维系统中最简单的2信道可用的例子。这里每个小区初始只分配了1个信道。前面提到这种情况有14种可能状态。将可用的两个信道记为1和2。将信道1分配给小区 i，信道2分配给小区 $i-1$ 和 $i+1$（信道1和2分别对应着图4.2中的集合 a 和 b）。很明显，如果每个小区里都有一个呼叫正在进行，则任何到达小区的呼叫将被阻塞。这对应的状态向量为(1,1,1)。其他的阻塞状态可以这样确定：考虑 $x_1=0,1$ 和最大值2，找到对应引起阻塞的 x_2 和 x_3 的值。我们将找到的9种状态在表4.1中列出，每个后面都有简单的注释，每个小区中占用的信道在括号中都有注明。

表4.1 阻塞状态，小区 i，定向信道锁定的信道借用方法，每小区信道数 $m=1$

x_1，小区 i	x_2，小区 $i+1$	x_3，小区 $i-1$	
0	0	2(1,2)	小区 $i-1$ 从小区 i 借取信道1，即使小区 $i+1$ 在信道2上没有正进行的呼叫，小区 i 也不能占用它
0	2(1,2)	0	由于对称性，同上
0	2(1,2)	2(1,2)	小区 $i-1$ 和小区 $i+1$ 每个都有 $2m=2$ 个呼叫在通话中，因而不违背复用准则
0	1(2)	2(1,2)	小区 $i-1$ 从小区 i 借取信道1
0	2(1,2)	1(1)	由于对称性，同上
1(1)	0	1(2)	小区 i 不能从小区 $i+1$ 借取未使用的信道2
1(1)	1(2)	0	由于对称性，同上
1(1)	1(2)	1(2)	在文中讨论
2(1,2)	0	0	小区 i 从其中一个邻近小区借用信道2，锁定另一个相邻小区的信道使用

我们将处在这些状态的概率累加，得到小区 i 的阻塞概率 B_i。更一般地，对于总共拥有 $2m$ 个信道，并采用了一维定向信道锁定的信道借用方法的小区 i，其阻塞概率由 Yeung 和 Yum(1995)给出

$$B_i=\sum_{x_1+x_3\leqslant 2m}P[x_1,2m-x_1,x_3]+\sum_{x_1+x_2<2m}P[x_1,x_2,2m-x_1] \tag{4.1}$$

这里的 $P[x_1,x_2,x_3]$ 是系统处于状态 (x_1,x_2,x_3) 的概率（为简单起见，不再使用状态向量这个名称，而只使用状态。从上下文明显可知，说的是1个小区还是3个小区的状态）。式(4.1)中各状态的累加正好是表4.1中所标记出的，这留给读者自行证明。

如何计算式(4.1)中的概率 $P[x_1,x_2,x_3]$ 来确定阻塞概率 B_i 呢？同第3章中固定信道的分配所考虑的一样，假定小区 i 中的呼叫按泊松速率 λ_i 到达，小区 i 的厄兰负载由 $A_i=\lambda_i/\mu_i$ 给出。类似地，认为小区 $i+1$ 和 $i-1$ 中的呼叫分别按泊松速率 λ_{i+1} 和 λ_{i-1} 到达。小区 $i+1$ 和 $i-1$ 中的呼叫持续时间则分别服从均值为 $1/\mu_{i+1}$ 和 $1/\mu_{i-1}$ 的负指数分布。假定业务性质都相同，则其呼叫持续时间均值都为 $1/\mu$。在泊松呼叫到达——指数保持时间的假设下，可以定

义图 4.3 中的多种状态间的转移关系。例如，在 3 小区系统中，i，$(i+1)$和$(i-1)$处在状态(x_1, x_2, x_3)下。比如，小区 i 中有一个呼叫到达(产生了一个请求)，如果这个呼叫能够被接纳，则状态转移到(x_1+1, x_2, x_3)。如果小区 $i+1$ 中有一个呼叫到达，则状态转移到(x_1, x_2+1, x_3)。相反地，如果小区 i 中有一个呼叫结束，状态转移到(x_1-1, x_2, x_3)。前面已经说明了小区 i 中的呼叫到达(请求)率为 λ_i，呼叫结束率为 μ。根据这些呼叫到达率和结束率，可以建立一张说明不同状态间转换速率的状态转移图。这样一张解耦的 3 小区 i，$(i+1)$和$(i-1)$的状态转移图见图 4.4。可以利用该图计算出所处状态(x_1, x_2, x_3)和稳态概率 $P[x_1, x_2, x_3]$，从而根据式(4.1)就能确定小区 i 的阻塞概率 B_i。值得注意的是，在图 4.4 所示小区 i 的邻近小区中，呼叫到达速率 λ_{i+1} 和 λ_{i-1} 比实际非解耦系统要小，这是由于小区 i 实际上可以使用一些在解耦系统中被锁定的信道。例如，信道 k 被小区 $i+1$ 借到小区 $i+2$，这并不造成对小区 i 的同信道干扰。如果这样，小区 $i+1$ 就不能使用该信道，但小区 i 可以使用。这样，就降低了小区 i 的阻塞概率。可利用式(4.1)计算得到的 B_i，它是实际阻塞概率的上界。

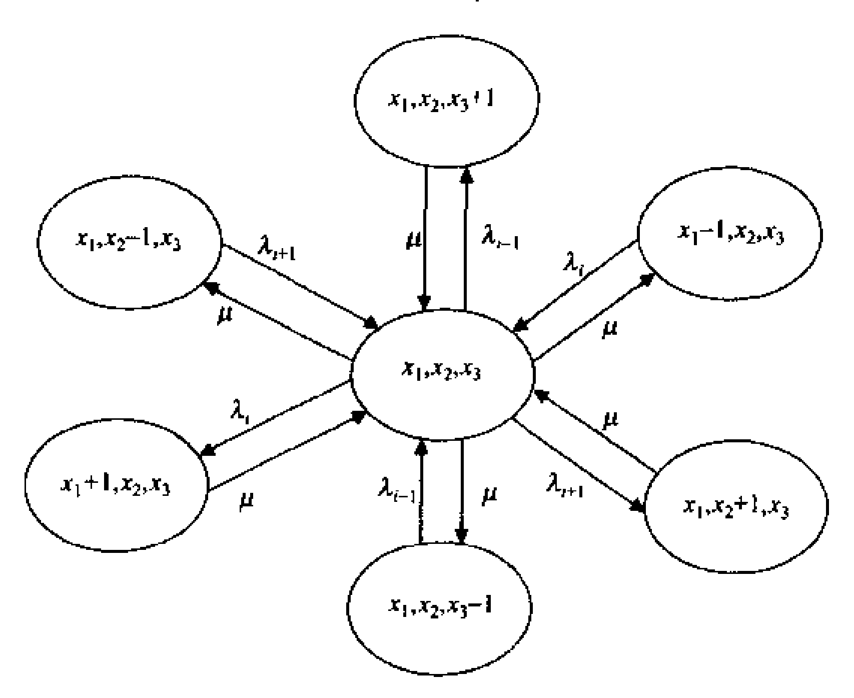


图 4.4　3 个解耦小区的状态转移图(引自 Yeung 和 Yum，1995：Fig.3)

针对蜂窝环境下的呼叫业务，对其做泊松呼叫到达——指数保持时间的假设(也是电话业务的通用模型)，以图 4.3 和图 4.4 为基础的状态转移图，描述了一种特殊形式的马尔可夫链的转换关系。这类课题多年来已被充分研究。特别在这个例子里，可以证明(Yeung 和 Yum，1995)理想的稳态概率 $P[x_1, x_2, x_3]$为 product-form 型(Schwartz，1987)[①]

$$P[x_1, x_2, x_3] = G^{-1} \prod_{j=1}^{3} A_j^{x_j} / x_j!, \quad A_j \equiv \lambda_j / \mu_j \tag{4.2}$$

术语 product-form 用来表明概率为若干项乘积，在此例中有 3 项。因此，即使 i，$(i+1)$和$(i-1)$3 个小区分别占用 x_1, x_2, x_3 个信道的状态明显是耦合的，其占用信道的状态概率和独

① Schwartz，M. 1987. *Telecommunication Networks: Protocols, Modeling, and Analysis*, Reading, MA, Addison-Wesley.

立占用的效果是一样的。它们并非是真正的独立,在式(4.2)中的 G^{-1} 项将其耦合,该项是必须的,以确保概率归一化。这样,G 项可以通过对 $P[x_1, x_2, x_3]$ 所有可能的状态归一化得到

$$\sum_{\text{所有状态}} P[x_1, x_2, x_3] = 1 \tag{4.3}$$

因而很明显,式(4.2)中的 G 项为

$$G = \sum_{\text{所有状态}} \left(\prod_{j=1}^{3} A_j^{x_j} / x_j! \right) \tag{4.4}$$

为了表明乘积形式和真正独立概率的解的不同,考虑对3个小区均采用固定信道分配(FCA)的情况。在这种情况中,有显式解

$$P[x_1, x_2, x_3] = P(x_1)P(x_2)P(x_3) \tag{4.5}$$

其中,

$$P(x_j) = \frac{A_j^{x_j}}{x_j!} \Bigg/ \sum_{x_j=1}^{m} \frac{A_j^{x_j}}{x_j!} \tag{4.6}$$

式(4.6)可很容易地从泊松呼叫到达——由指数保持时间的模型的基础排队论得到(Schwartz, 1987)。很容易证明,对 x_j 所有的值进行累加,式(4.6)刚好归一化。注意,在 $x_j = m$ 时,可以得到厄兰-B阻塞公式,也就是第3章中的式(3.7)。暂时用式(3.7)对动态信道分配和固定信道分配进行比较(第9章中有更多的关于泊松呼叫到达和厄兰公式的讨论)。

这一章集中讨论的是一维的采用定向信道锁定的信道借用的动态信道分配策略的例子,由式(4.2)和式(4.4),可以计算该例中小区 i 的阻塞概率 B_i(回忆一下评述,式(4.2)和式(4.3)提供了 B_i 的上界)。给定 B_i,整个 N 小区系统的阻塞概率 B 由加权和给出

$$B = \sum_{i=1}^{N} \lambda_i B_i \Bigg/ \sum_{i=1}^{N} \lambda_i \tag{4.7}$$

在遇到特性相同的情况时,如业务均匀的例子,式(4.1)中的 $\lambda_i = \lambda, \mu_i = \mu, B_i = B$。考虑两种特殊情况:$m = 1$(系统中有 $2m = 2$ 个信道可用)和 $m = 2(2m = 4)$(这里采用的例子主要出自参考文献——Yeung 和 Yum, 1995,该书中有大量的例子可供参考)。对于特性相同的情况,每个小区中的业务到达率 λ 相同,则式(4.2)中的厄兰负载 A_j 都等于 A。首先考虑每个小区分配 $m = 1$ 个信道的情况。将式(4.2)带入式(4.1),容易证明,阻塞概率 B 有如下表达式:

$$B = \frac{\frac{7}{2}A^2 + 2A^3 + \frac{A^4}{4}}{1 + 3A + \frac{9}{2}A^2 + 2A^3 + \frac{A^4}{4}} \tag{4.8}$$

将式(4.2)中的分子乘积项按表4.1中的9种阻塞状态相加,化简和的结果,可得到式(4.8)中的分子。分母则对应式(4.2)中 G 的计算:9种阻塞状态和5种非阻塞状态的乘积的和。将式(4.8)作为厄兰负载 A 的函数,作图见图4.5。在同一张图上做出等效的2信道($m = 1$)的固定信道分配方案的阻塞概率性能曲线,以作为对比。在固定信道分配方案中,相邻小区交替分配两信道中的一个。本例中平均每个小区有 $m = 1$ 个信道,阻塞概率可简单表示为

$$B = A/(1 + A) \tag{4.9}$$

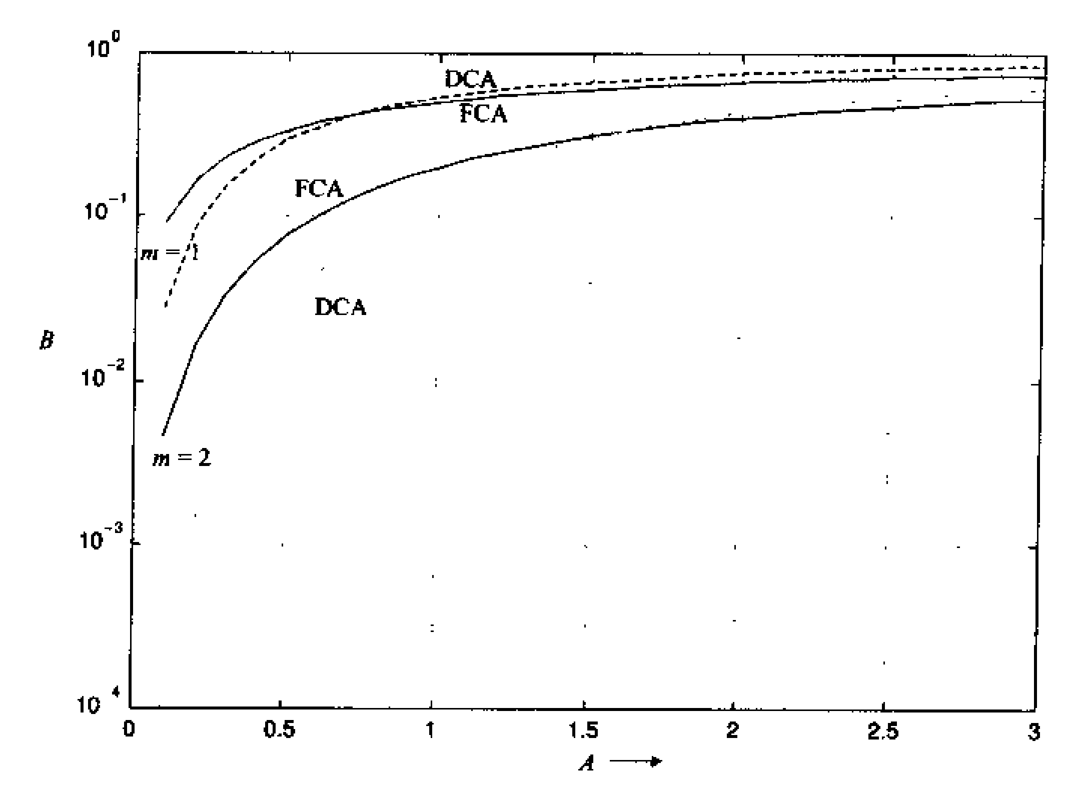


图 4.5　一维定向信道锁定的信道借用方案与固定信道分配的阻塞概率的比较

利用式(3.7)或式(4.6)，其中 $x_j = m = 1$。注意到在这种可能的最简单的例子中，固定信道分配在负载小于 $A = 0.75$ 的情况下，其阻塞概率性能比固定信道分配有所改善。在大于这个负载的情况下，正如所预料的，固定信道分配性能稍好一些。图 4.5 也画出了平均每个小区有 $m = 2$个信道，也就是系统总共有 4 个信道的情况。请读者自行通过式(4.1)至式(4.4)，计算动态信道分配的具体结果。回忆前面对图 4.3 的讨论，在线性一维的定向信道锁定的信道借用策略的第 2 个例子中，共有 55 个状态。在这 55 个状态中，有 25 个为阻塞状态，其他 30 个为非阻塞状态。在本例中，类似于式(4.8)的表达式可计算阻塞概率，其同样也包含对式(4.2)中阻塞状态的分子乘积项求和，以及对分母中出现的乘积项按 55 个状态求和。这个例子的最终性能曲线，按阻塞概率 B 对应负载 A 作图，如图 4.5 所示。对等效的 $m = 2$ 的固定信道分配的情况也同时作图。该曲线的表达式为

$$B = A^2/2/(1 + A + A^2/2) \tag{4.10}$$

该表达式同样可以通过令式(3.7)或式(4.6)中 $x_j = m = 2$ 得到。注意到在本例中，在厄兰负载小于 $A = 2$ 时，固定信道分配方法比等效的固定信道分配方法有较大的性能改善。大于这个负载的固定信道分配能提供较好的性能。前面解释过在中低业务负载的情况下，动态信道分配相对固定信道分配在性能上有所改善，而在重负荷下，固定信道分配的性能却较好。这些结果是吻合的。

由图 4.5 中所考虑的最简单的两个例子，我们注意到，当信道数目增加时，动态信道分配应用的负载的变化范围也增大了。这已经被 Yeung 和 Yum(1995)计算的例子所证明，文中考虑了 $2m = 20$ 个信道。该论文中假定额定负载为 $\lambda = 100$ 呼叫/h。平均呼叫持续时间 $1/\mu = 3$ min，可得厄兰负载为 $A = 5$ 厄兰。在这个负载值下，固定信道分配的阻塞概率为 0.12。定向信道锁定的信道借用策略能够达到约小于一个数量级的阻塞概率。两种方案的差距在超过额定负载 50%时仍维持不变。在负载相当重的情况下，两种方案估计会接近，并且最终固定信

道分配会有优势。同时,正如我们所预料的,在系统中有20个可用信道的情况下,与前面只考虑2个或4个可用信道的情况相比,可调整和容纳的负载要大很多。

如前所述,为简化起见,我们集中讨论了一种动态信道分配方案,而避免了陷入到各种各样动态信道分配策略的细节解释中。利用这种动态信道分配方法作为例子,我们也证明了利用动态信道分配是如何改善固定信道分配在中低负荷区域的阻塞概率的。特别地,提出将动态信道分配引入GSM带来实质性的性能改善已经被仿真证明(Priscoll等,1997[①];Le Bris和Robison,1999[②])。第1章中提到过GSM系统是世界范围内成功采用的3种第二代蜂窝网络之一。GSM将在第6章和第8章中描述。Priscoll等在1997中的仿真研究表明,引入动态信道分配后,相对于固定信道分配,在阻塞概率为1%的范围内,GSM网络可以多容纳大约40%~50%的用户。Le Bris和Robison(1999)总结的研究针对的是70%负载的情况,也就是在70%的信道忙时,各用户在瑞利衰落的环境下的差错率。他们的结果表明,利用动态信道分配在上行方向可以将差错率改善25%,下行可改善15%(差错率的概率在第6章有简要总结)。

其他由于采用动态信道分配获得改善的例子在Katzela,Naghshineh(1999)的著作和书后的参考文献中出现过。更一般地,利用定界技术,通过研究能提供最优阻塞概率性能的理想形式的动态信道分配,动态信道分配技术的优点已经得到证明。这种理想形式的动态信道分配称为最大紧致策略,Everitt和Manfield(1989)[③]通过蒙特卡罗仿真,Raymond(1991)[④]从理论上针对一些简单的蜂窝结构估计了其性能。

在最大紧致策略中,新到达的呼叫总能够得到一个可用信道,可能必须通过重新调整正在进行的呼叫的信道分配。惟一的约束就是复用限制必须满足。Everitt和Manfield用他们的模型在350个信道,49个六边形小区,7小区复用的条件下比较了最大紧致策略和固定信道分配。在固定信道分配中每个小区占用50个信道。当业务负载在使用固定信道分配的情况下导致2%和5%的阻塞概率时,他们发现最大紧致策略,也就是动态信道分配,相对固定信道分配,其阻塞概率有一个数量级的下降。这个结果与前面引用的Yeung和Yum(1995)的结果符合。更明确地说,在他们的例子中,要保持阻塞率为2%,固定信道分配能容纳的平均业务负载为40.3厄兰/小区,而如果允许的阻塞率增加到5%,平均业务负载增加到44.5厄兰/小区。在相同的负载下,最大紧致策略阻塞概率低1个数量级。作为另一种选择,最大紧致策略或动态信道分配在阻塞率为2%时,平均业务负载为47.3厄兰/小区;在阻塞率为5%时,负载为50.7厄兰/小区。这样,通过采用最大紧致方法可容纳的业务增加了15%甚至更多。

Everitt和Manfield也研究了业务波动对最大紧致和动态信道分配的相对性能的影响。他们把每个小区内的业务建模为均值,为前面提到的数值的正态分布的随机变量增大方差,以模仿其增加的波动性。得到的结果表明,在动态信道分配下,相对于固定信道分配,其阻塞概率

---

① Priscoll, F. Della *et al*. 1997. "Application of dynamic channel allocation strategies to the GSM cellular network," *IEEE Journal on Selected Areas in Communications*, 15, 8(October), 1558 - 1567.

② Le Bris, L. and W. Robison. 1999. "Dynamic channel assignment in GSM networks," Proc. VTC'99, IEEE Vehicular Technology Conference, Amsterdam, September, 2339 - 2342.

③ Everitt, D. and D. Manfield. 1989. "Performance analysis of cellular mobile communication systems with dynamic channel assignment," *IEEE Journal on Selected Areas in Communications*, 7, 8 (October), 1172 - 1180.

④ Raymond, P.-A. 1991. "Performance analysis of cellular networks," *IEEE Transactions on Communications*, 37, 12(December), 1787 - 1793.

对于业务波动性的敏感程度要低很多。Raymond 利用由线性小区模型得到的理论性能结果，与业务均匀的情况相比是相似的。他还发现在利用固定信道分配会带来1%～2%的阻塞概率的业务负荷下，采用最大紧致/动态信道分配方法能使阻塞概率降低约1个数量级。当平均流量下降时，对这种阻塞概率性能的改善会减小。在流量为3倍的平均负载情况下，动态信道分配和固定信道分配性可相同，阻塞概率均上升为40%左右。这个值显然对应的是高负载的情况。两种对这种理想动态信道分配的研究都证实了在一些动态信道分配特例中的结果，其中包括本节中讨论的定向信道锁定的信道借用方法：在标称负荷下，采用固定信道分配会导致1%～5%的阻塞概率，采用动态信道分配则能获得明显的性能改善。采用动态信道分配还能够极大地降低业务波动性造成的影响。然而，这种性能的改善随着负荷的增加在减小，直到高负荷时改善消失，动态信道分配则在足够高负荷情况下占优势。

4.2　功率控制

在介绍移动无线通信系统中的小区概念时，注意到这个概念是由增大系统容量的需要引发的。如果使用相同信道的小区不是“特别接近”，就能使更多的用户接入系统，系统信道从而得以复用。“接近度”的规定保证来自同信道小区的总干扰功率不大于本小区内信道功率的一个确定比值。通过对一给定信道引入信干比 SIR 来量化这个概念，以作为性能的度量。给定了 SIR 特定的要求和总的可用信道数目，就可以确定任何一个给定小区中能够容纳的移动台用户数目(从而得到整个移动系统)。每个特定用户都期望获得一定的阻塞概率作为性能目标，可以通过确定这些指标来实现上述计算。正如上节中提到的，使用动态信道分配能使系统在适度的业务负载和给定的信道数目条件下，比固定信道分配容纳更多的用户。

所有因素中最关键的是保持合适的 SIR。这需要监测和控制移动功率以保证在一个给定小区内不低于一个确定的阈值。所有的移动蜂窝系统采用了在基站检测上行功率，再对其小区内的移动单元发送功率控制信息的机制。在第6章、第8章和第10章对码分多址系统的讨论中，将看到严格的功率控制对于其系统的正常工作是至关重要的。在本节中，将结合和比较文献中提出的两种分布式功率控制策略，其中，对发送功率进行独立调整，使得所有接收机的 SIR 尽可能大。我们将会证明最终导致所有接收机的 SIR 相同。这些在文献中提到的方法及其改进形式主要用于有复用限制的系统，如第3章中讨论的。这些系统以第6章和第8章(D-AMPS 或 IS-136 和 GSM)中讨论的第二代 TDMA/FDMA 系统(时分多址/频分多址)为典型(时分多址和频分多址的概念在第6章中讨论)。将在第6章和第8章讨论的第二代 CDMA(码分多址)系统 IS-95，还有第10章中将讨论的第三代码分多址系统，都没有复用限制：每个小区可能使用相同的频率。我们将在第6章看到，需要利用编码来处理干扰信号。本节中讨论的两种方法以及用来比较它们的例子，对码分多址系统是不适用的。我们会效仿这些方法的细节讨论，通过简单介绍一种用于码分多址系统中控制发送功率的方法，改进得到这些及类似方法的一般形式。我们会证明这种方法确实和最先讨论的时分多址/频分多址中的方法有共同点。

为简化各种方法的讨论，集中讨论上行功率控制。假设发射机位于移动台，接收机位于基站。这样每个基站对给定信道的 SIR 值进行测量。这种情况实际上在干扰台使用同样的传输频带的码分多址系统中更为关键。在描述这些方法的时候，会明白它们同样适用于下行功率控制，下行功率控制能保证在某个给定的移动终端测得的 SIR 足够大。

回忆3.2节和3.3节，上行SIR是给定基站通过某一信道接收到某一移动台的信号功率与在同一基站通过同一信道接收到其他移动台的干扰功率的比值。如在第3章中的分析一样，我们只着重讨论平均功率，通过将小区 i 内的 SIR_i 写做如下形式，可进行一般化讨论：

$$SIR_i \equiv \frac{P_i/d_{ii}^n}{\sum_{j=1,j\neq i}^{M} P_j/d_{ij}^n} \tag{4.11}$$

如图4.6所示，d_{ii} 是在小区 i 中使用所考察的信道的移动台到基站的距离，其发送功率为 P_i；d_{ij} 是在小区 j 中使用相同信道的移动台到小区 i 基站的距离，P_j 为其发送功率。参数 n 是第2章中讨论的路径衰减因子，典型值在2到4间变化。这里假设有 M 个使用相同频率的干扰移动台，每个位于不同的小区中。式(4.11)可被改写为另一种更为一般的形式

$$SIR_i = \frac{P_i G_{ii}}{\sum_{j=1}^{M} P_j G_{ij} - P_i G_{ii}} = \frac{P_i}{\sum_{j=1}^{M} P_j (G_{ij}/G_{ii}) - P_i} \tag{4.12}$$

G_{ij} 项成为从发射机 j(上行的第 j 个移动台)到接收机 i(第 i 个基站)的所谓的增益因子(实际上是损耗!)。在这里讨论的平均功率的情况下，当然有 $G_{ij} \equiv 1/d_{ij}^n$。功率控制的目标是为了保证系统中每个SIR都高于一个必须的阈值

$$SIR_i \equiv \gamma_i \geqslant \gamma_0, 1 \leqslant i \leqslant M \tag{4.13}$$

为了简化表示，今后如式(4.13)中一样，利用 γ_i 来表示 SIR_i。

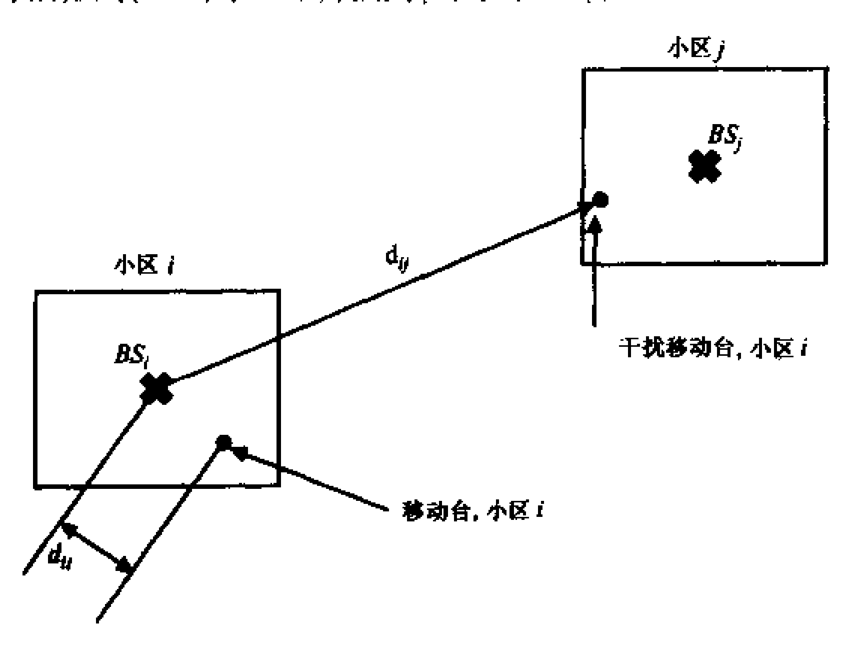


图4.6 计算 SIR_i 涉及到的距离

一些关于功率控制课题的论文都出现了式(4.13)，它表明需要将每个SIR都保持在阈值之上，我们将引用一部分。读者可以查阅每篇论文的参考书目以获得详尽的文献清单。特别地，分布式方法首先应用于卫星系统的功率控制。20世纪70年代出现的由 Aein(1973)[①]和

① Aein, J. M. 1973. "Power balancing in systems employing frequency reuse," *Comsat Technology Review*, 3, 2(Fall).

Meyerhoff(1974)[①]撰写的论文,是最早描述这类方法的论文。瑞典的 Zander 和他的小组是无线蜂窝系统功率控制算法的先驱。这里主要集中讨论由 Zander(1992)[②]提出的一种分布式功率控制方法,即分布式平衡算法或 DBA。这里的目标是在 M 个移动台中,在每个与各自的基站协同工作的情况下,用一种分布式的方式调整发送功率的值 P_i, $1 \leq i \leq M$,以获得最大可达到的 SIR。这个量被定义为在所有移动台功率最小 SIR 的取值中的最大值。如上面提到的,利用符号 γ 来表示 SIR,得到最大 SIR

$$\gamma^* = \max_{allpowers} \min \gamma_i \qquad 1 \leq i \leq M \tag{4.13a}$$

Zander 提出的分布式平衡算法,以及在这段讨论后介绍的另一种方法,都能够让使用相同信道的全部 M 个移动台均达到相同的最大 SIR 值 γ^*。可见,使用这种方法要付出的代价,必然是每个移动台和其基站间以及 M 个基站间的通信量受到限制(稍后将讨论的这种方法和相关方法的改进算法 Foschini 和 Miljanic(1993)[③],该算法可消除该问题)。分布式平衡算法是一种逐步迭代算法,需要输入整个增益集, $G_{ij} = 1/d_{ij}$, $1 \leq i, j \leq M$。它还假定了方法运作时这些增益不变。这也意味着快速的收敛到需要的 SIR 值 γ^* 是必须的。可惜的是,这种方法被证明需要相对长的时间收敛。第 2 种方法,也就是这种方法的改进形式,收敛则快很多。

两种方法都需要从初始功率值的集合开始,每个移动台都同时反复地调整各自的功率,来进行分布迭代。两种方法的调整策略不同。考虑第 i 个移动台的第 n 次迭代,为了简化表示,令 $r_{ij} \equiv G_{ij}/G_{ii}$,也就是用 G_{ii} 将 G_{ij} 归一化。由分布式平衡算法进行的功率调整如下:

$$P_i^{(n)} = c^{(n-1)} P_i^{(n-1)} \left[1 + \frac{1}{\gamma_i^{(n-1)}} \right] \tag{4.14}$$

其中,

$$\gamma_i^{(n-1)} = P_i^{(n-1)} \Bigg/ \left(\sum_{j=1}^{M} P_j^{(n-1)} r_{ij} - P_i^{(n-1)} \right) \tag{4.15}$$

常数 $c^{(n-1)}$ 必须比 0 大,并且这个值由所有移动台决定。然后,利用式(4.15)将式(4.14)化简,得到

$$P_i^{(n)} = c^{(n-1)} \sum_{j=1}^{M} P_j^{(n-1)} r_{ij} \tag{4.16}$$

式(4.16)描述的功率迭代被使用同一信道的每个移动台同时执行,也可以更简单地写成向量形式。这样,定义 $\boldsymbol{P}^{(n)}$ 为一个具有 M 个元素的向量,向量的第 i 项为 $P_i^{(n)}$。定义 $\boldsymbol{R}$ 为 $M \times M$ 矩阵,其第(i,j)项为 r_{ij}。注意到 $\boldsymbol{R}$ 为非负矩阵,对角项为 1,非对角项大于 0 但小于 1,这是因为 $r_{ii} \equiv G_{ij}/G_{ii} = 1$ 和 $r_{ij} \equiv G_{ij}/G_{ii} < 1$, $i \neq j$,则式(4.16)的分布式平衡算法可以简单地写成向量形式

---

① Meyerhoff, H. J. 1974. "Method for computing the optimum power balance in multibeam satellites," *Comsat Technology Review*, 4, 1 (Spring).

② Zander, J. 1992. "Distributed cochannel interference control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, 41, 3(August), 304 – 311.

③ Foschini, G. J. and Z. Miljanic. 1993. "A simple distributed autonomous power control algorithm and its convergence," *IEEE Transactions on Vehicular Technology*, 42, 4 (November), 641 – 646.

$$P^{(n)} = c^{(n-1)}RP^{(n-1)} \tag{4.17}$$

由于 R 为对角项为 1 的非负矩阵，利用非负矩阵的 Perron-Frobenius 定理(例如，可参考 Noble 和 Daniel，1988[①])，可以很快证明其最大的特征值 $z^* > 1$，且重数为 1，其他的所有 $M-1$ 个特征值都满足 $|z_j| < z^*$（R 为 $M \times M$ 矩阵，因此有 M 个特征值）。特别地，现在将证明式(4.17)中的分布式平衡算法会让每个小区的 SIR 收敛到式(4.18)给出的相同极值 γ^*。

$$\lim_{n\to\infty} \gamma_i^{(n)} = \gamma^* = 1/(z^* - 1) \tag{4.18}$$

很明显，要让 SIR 值 γ^* 尽可能大，就得让特征值 z^* 尽可能接近 1。但稍后会看到，如果 z^* 越接近 1，这种方法收敛得越慢。这一点引发了第 2 种方法的讨论，它收敛到相同的 SIR 值 γ^*，但收敛得更快。

下面是分布式平衡算法的证明，该方法同时应用于使用相同信道且相互之间产生干扰的所有 M 个移动台，会使所有移动台的 γ_i 收敛到相同值 γ^*：让每个移动台独立选择初始发送功率。这就建立了初始发送功率向量 $P^{(0)}$。我们将(4.17)迭代 n 次，得到

$$P^{(n)} = R^n \left[\prod_{j=0}^{n-1} c^{(j)}\right] P^{(0)} \tag{4.19}$$

在上文注意到 $M \times M$ 矩阵 R 有 M 个特征值 $z_j, 1 \leqslant j \leqslant M$。每个都有其对应的特征向量 ϕ_j。这些 M 个特征向量代表了一个正交集，它能够作为一组基对初始功率向量 $P^{(0)}$ 做惟一展开

$$P^{(0)} = \sum_i a_i \phi_i$$

式(4.20)中出现的展开系数 a_j 由常见的各自的点积给出

$$a_j = (P^{(0)}, \phi_j)/(\phi_j, \phi_j) \tag{4.21}$$

注意，若 x 和 y 每个都是 M 维向量，$(x, y) \equiv \sum_j x_j y_j$。

利用式(4.20)、式(4.19)中的 $R^n P^{(0)}$，可由 $\sum_j a_j R^n \phi_j$ 替换，但是注意到，通过定义特征值和特征向量，$R\phi_j = z_j \phi_j$。因此，$R^n \phi_j = z_j^n \phi_j$ 且 $R^n P^{(0)} = \sum_j a_j z_j^n \phi_j$。为了不失一般性，假设 z_1 为 $z^* > 1$，M 个特征值中最大的一个，则对于足够大的变量 n 次迭代，$(z^*)^n$ 肯定能在对 j 求和之后占优，且 $\sum_j a_j z_j^n \phi_j \to a_1 (z^*)^n P^*$。这里用符号 P^* 代表 $z_1 \equiv z^*$ 对应的特征向量 ϕ_1，因而 $\lim_{n\to\infty} R^n P^{(0)} = a_1 (z^*)^n P^*$，而且由式(4.19)可得

$$\lim_{n\to\infty} P^{(n)} = P^* a_1 \lim_{n\to\infty} (z^*)^n \prod_{j=0}^{n-1} c^{(j)} \tag{4.22}$$

现在的目标是在第 j 次迭代时选择常数 $c^{(j)}$，使得式(4.22)的极限有界，也就是防止 M 个移动台的功率变得过大。一种可能的方法是简单地让 $c^{(j)} = 1/\max(P^{(j)})$，这里利用 $\max(P^{(j)})$ 表示在第 j 次迭代时 M 个发送功率中的最大值。但注意到，这种选择需要所有移动台在执行这种方法的第 j 次迭代时相互之间同步地交换各自的功率值。另一种选择是每个基站与其他基站交换它的移动台功率值，再传给每个小区内的移动台。

简要说明一下现在的结果，已经证明了分布式平衡算法中，在第 n 次迭代功率向量 $P^{(n)}$ 的

① Noble, B. and J. H. Daniel. 1988. *Applied Linear Algebra*, 2nd edn, Englewood Cliffs, NJ, Prentice-Hall.

第 i 个分量,也就是第 i 个小区中的发送功率,当 n 足够大时,逼近下面的表达式

$$P_i^{(n)} = c^{(n-1)} P_i^* a_1 (z^*)^n \prod_{j=0}^{n-2} c^{(j)} \tag{4.23}$$

注意到 $P_i^{(n)}$ 也可以写成式(4.14)的形式。将式(4.14)中的 $P_i^{(n-1)}$ 用等效的式(4.23)替换,将 n 换为 $n-1$,也能得到

$$P_i^{(n)} = c^{(n-1)} \left[1 + \frac{1}{\gamma_i^{(n-1)}} \right] P_i^* a_1 (z^*)^{n-1} \prod_{j=0}^{n-2} c^{(j)} \tag{4.24}$$

比较式(4.23)和式(4.24),对 n 取极限,则显然有

$$\lim_{n\to\infty} \gamma_i^{(n-1)} \approx \gamma^* = 1/(z^* - 1) \tag{4.25}$$

这个表达式就是式(4.18)提到过的,所有 M 个移动台的 SIR 值都会收敛到共同值。

现在考虑这种方法在两个一维例子中的应用。第 1 个例子是图 4.7 所示的 11 个小区的线性蜂窝系统。假定信道在每两个小区间复用。着重关心斜线标出的 6 个小区。基站用空心圆表示;6 个可能相互干扰的移动台用黑点表示,每个都分配相同的信道。移动台的位置全部选在小区边界,对应最差情况下的干扰(注意到距离 d_{ij} 和对应的 SIR 值完全取决于不同移动台相对基站的位置。当移动台移动时,距离会变化。这就是上面提到的为什么需要找到一种快速收敛的功率控制方案的原因)。每个小区都规定占有一个单位的宽度(也可以用其他的距离度量,将宽度规定为 D 个单位,但在处理不同距离的功率比时,计算矩阵 $\boldsymbol{R}$ 中的元素 $r_{ij} = G_{ij}/G_{ii} = d_{ii}^n/d_{ij}^n$,故 D 的具体值可消去)。留给读者证明这个例子中的矩阵 $\boldsymbol{R}$ 为式(4.26)中的形式

$$\boldsymbol{R} = \begin{bmatrix} 1 & \frac{1}{3^4} & \frac{1}{7^4} & \frac{1}{11^4} & \frac{1}{15^4} & \frac{1}{19^4} \\ \frac{1}{3^4} & 1 & \frac{1}{3^4} & \frac{1}{7^4} & \frac{1}{11^4} & \frac{1}{15^4} \\ \frac{1}{7^4} & \frac{1}{5^4} & 1 & \frac{1}{3^4} & \frac{1}{7^4} & \frac{1}{11^4} \\ \frac{1}{11^4} & \frac{1}{9^4} & \frac{1}{5^4} & 1 & \frac{1}{3^4} & \frac{1}{7^4} \\ \frac{1}{15^4} & \frac{1}{13^4} & \frac{1}{9^4} & \frac{1}{5^4} & 1 & \frac{1}{3^4} \\ \frac{1}{19^4} & \frac{1}{17^4} & \frac{1}{13^4} & \frac{1}{9^4} & \frac{1}{5^4} & 1 \end{bmatrix} \tag{4.26}$$

利用数学软件,很容易得到对应的 6 个特征值为 1.0137,1.006,1,0.999,0.994,0.987。对应的最大特征值为 $z^* = 1.0137$,所有移动台都收敛到 SIR 值为 $\gamma^* = 1/(z^* - 1) = 73.3$,或以分贝为单位即 18.7 dB。

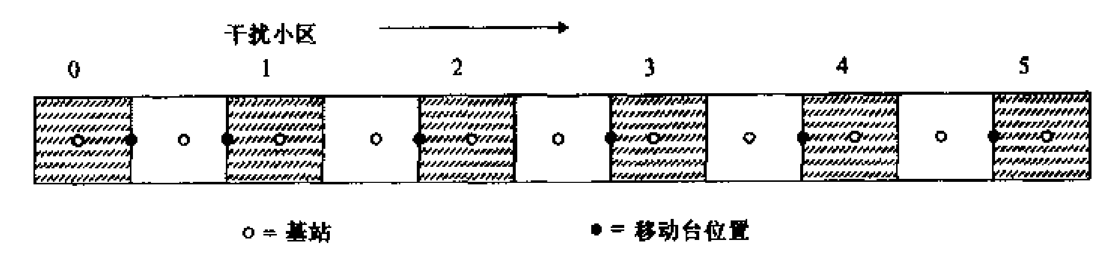


图 4.7　11 个小区的线性蜂窝系统

作为第2个例子，只改变其中两个移动台的位置，在图4.8中构造出对称的11个小区的情况。注意到这个简单变化在功率控制上产生了影响。这个例子中的 $\boldsymbol{R}$ 是对称的。很容易证明矩阵 $\boldsymbol{R}$ 的6个特征值为1.0140，1.012，1.011，0.988，0.988，0.987，这样，$z^* = 1.0140$ 且 $\gamma^* = 71.6$ 或者18.5 dB。注意，由于移动的引入，会造成SIR的微小变化。移动台位置的显著变化会引起相应的SIR更大的变化。

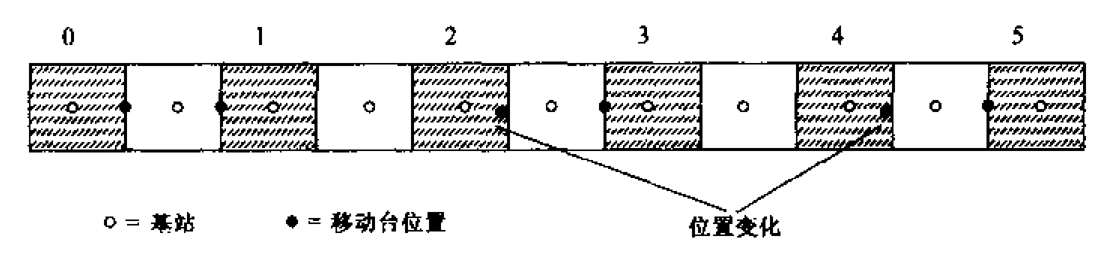


图4.8 移动台位置对称的情况

由于距离和增益的参数变化，功率控制方法收敛到的SIR取决于移动台的位置，正如前面所说，算法的快速收敛是很重要的(回忆一下前面的假设，所有的增益参数和它们之间的距离，都是不变的)。但前面也注意到所关心的分布式平衡算法收敛较慢。从刚才介绍的例子可以清楚地看到原因——特征值很接近。但在证明收敛性时曾经假设，对于足够大的 n 次迭代 $z_1^n \gg z_i^n, i \neq 1$。然而由于 z_i 和 z_1 接近，在这种条件下 n 必须非常大才能成立。注意到两个例子中的特征值都很接近，所以其收敛速率估计会很慢。利用简单的矩阵分析容易证明，这种方法总是收敛得很慢：矩阵 $\boldsymbol{R}$ 的特征值总是一堆在1附近的数，这个现象可以证明如下。矩阵 $\boldsymbol{R}$ 的迹 $\mathrm{Tr}(\boldsymbol{R})$ 定义为对角元素的和，也可由特征值的和得到。这样有

$$\mathrm{Tr}(\boldsymbol{R}) \equiv \sum_{i=1}^{M} r_{ii} = \sum_{i=1}^{M} z_i \tag{4.27}$$

例子中，对角元素 r_{ii} 全部等于1，$\mathrm{Tr}(\boldsymbol{R}) = M$ 值。如果想得到一个高的SIR值 γ^*，必然有 $z^* - 1 = z_1 - 1 \approx 1/\gamma^* \ll 1$，这样 $z^* = 1 + \varepsilon, \varepsilon \ll 1$(注意到在两个例子中，分别有 $\varepsilon = 0.0137$ 和 0.014)。由于 $\mathrm{Tr}(\boldsymbol{R}) = M$，由式(4.27)有

$$\sum_{j=2}^{M} z_j = M - z^* \approx M - 1 \tag{4.28}$$

这样大的 M 值意味着所有的 z_j 都接近于1且相隔很近，但是期望的收敛速率很慢。事实上，所参照的两个例子，即使 M 仅仅为6，我们也注意到6个特征值是多么接近。考虑图4.7的第1个例子中两个最大特征值的比值。这里有 $z_2/z_1 = 1.006/1.0137 \approx 0.992$。这样对于20次迭代，$z_2^n/z_1^n$ 仅仅是0.85，而非是定理收敛性推导中的假设 $\ll 1$。当 $n = 60$ 时，z_2^n/z_1^n 仍只有0.52。其他的特征值相对也很接近 $z^* = z_1$，直到 n 非常大时才开始收敛。

如果 z_1 和 z_2 都小于1会如何呢？那么收敛速率会快很多。这就引出了第2种方法，我们将会看到该方法的特征值都比现在的特征值小1；也就是说，如果将它们记做 $\lambda_i, 1 \leqslant i \leqslant M$，将证明 $\lambda_i = z_i - 1$。在图4.7第1个例子中，就有 $(\lambda_2/\lambda_1)^n = (0.006/0.0137)^n = 0.44^n$。对于 $n = 8$，这个值为0.0015，很明显收敛很紧。对于图4.8第2个例子，有 $(\lambda_2/\lambda_1)^n = 0.86^n$，收敛速率慢一些，

但仍比分布式平衡算法慢很多。第2种改进方法由 Grandhi, Vijayan 和 Goodman(1994)[①]提出，被简单称为分布式功率控制或者分布式功率控制算法，是基于卫星系统提出的 Meyerhoff 功率平衡方法(Meyerhoff,1994)。除了利用分布式平衡算法中的式(4.14)进行功率迭代之外，它还简单采用了迭代

$$P_i^{(n)} = c^{(n-1)} P_i^{(n-1)} / \gamma_i^{(n-1)} \tag{4.29}$$

利用矩阵向量形式可以写做

$$\boldsymbol{P}^{(n)} = c^{(n-1)} \boldsymbol{A} \boldsymbol{P}^{(n-1)} \tag{4.30}$$

其中，A 是一个 $M \times M$ 矩阵，其元素 $a_{ij} \equiv G_{ij}/G_{ii}, i \neq j; a_{ii} = 0$。除了减去一个单位阵 $\boldsymbol{I}: \boldsymbol{A} = \boldsymbol{R} - \boldsymbol{I}$，矩阵 $\boldsymbol{A}$ 与分布式平衡算法中的矩阵 $\boldsymbol{R}$ 是完全相同的。矩阵 $\boldsymbol{A}$ 的 M 个特征值 $\lambda_j, 1 \leqslant j \leqslant M$，可由 $\lambda_j = z_j - 1$ 简单给出，其中 z_j 为前面讨论的矩阵 $\boldsymbol{R}$ 的特征值。用上面证明分布式平衡算法收敛性的相同方法，容易证明分布式功率控制算法收敛到与分布式平衡算法相同的 SIR 值 γ^*，但由 $\gamma^* = 1/\lambda^*$ 给出，其中 $\lambda^* = z^* - 1$ 为矩阵 $\boldsymbol{A}$ 最大的特征值。收敛速率由 λ_2/λ^* 决定，它是一个比分布式平衡算法中决定收敛速率的 z_2/z^* 小很多的数。这点在上面已经说明。

考虑与前面图4.7和图4.8中讨论的11个小区线性网络相同的两个例子。图4.9给出了这两个例子中小区3的 SIR 值 $\gamma_3 = \text{SIR}_3$ 的收敛性。注意到图4.9(b)所示的例1的小区3的 SIR，和预计的一样，虽然初始波动得比较厉害，但在8次迭代之内确实收敛得很快。图4.8中对称情况下对应的 SIR 也跟估计的一样，在20次迭代中收敛得很慢，但是其波动更小。Grandhi 等(1994)对一维50个小区的系统做了类似的仿真。每3个小区复用系统中各个信道，结果自然得到更高的 SIR。他们选择了3种具有代表性的移动台位置，而不是在这些例子中选择这两种情况。他们发现，最多需要5次迭代就能达到最小 SIR 约为31.8 dB到34 dB。对应的分布式平衡算法的收敛速率慢很多，在考虑的3种情况中，经过50次迭代分布式平衡算法都远远没有真正的收敛。

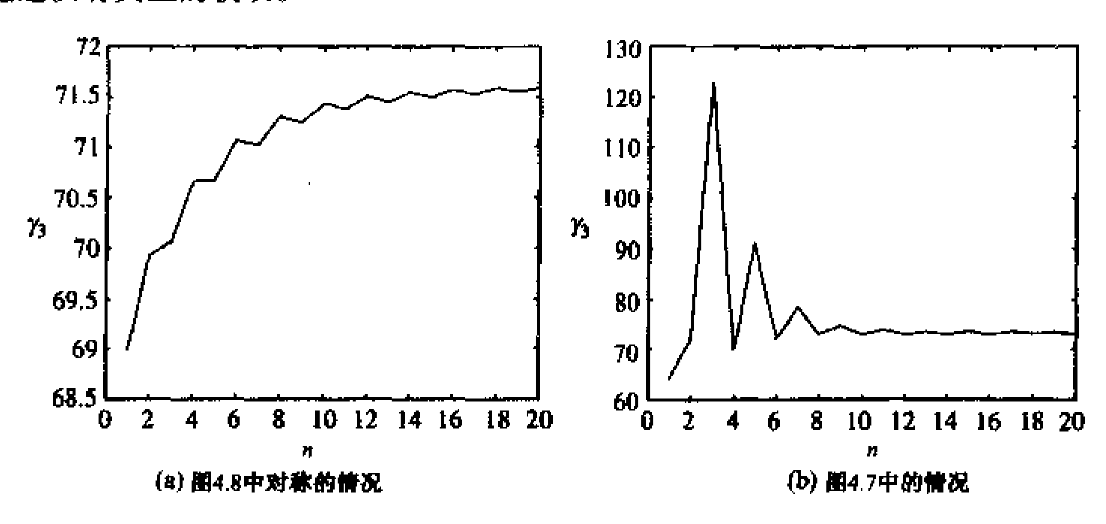


图4.9 图4.7和图4.8的例子中分布式功率控制算法的收敛性

① Grandhi, S. A. *et al*. 1994. "Distributed power control in cellular radio systems," *IEEE Transactions on Communications*, 42, 2/3/4 (February-April), 226–228.

到目前为止,我们还集中讨论了分布式功率控制中的 SIR 收敛性。当收敛完成时每个移动台的发送功率是多大呢?考虑 11 个小区的例子。仍假设每两个之间复用,6 个小区分配相同信道。在图 4.7 的第 1 种情况下,相对于功率最大的那个,每个移动台的相对功率可由下面的列表给出:

移动台 →	0	1	2	3	4	5
相对功率 →	0.86	1	0.214	0.052	0.016	0.004

注意,这种方法导致移动台 3,4,5 的发送功率相对较小。考虑第 2 种情况,由于 6 个移动台中的两个移动了它们在情况 1 中的位置,对应的相对功率列表如下:

移动台 →	0	1	2	3	4	5
相对功率 →	0.64	0.71	1	1	0.71	0.64

由于小区 2 左侧和小区 3 右侧完全相同,根据对称性可以预料到这个结果。

蜂窝系统分布式功率控制的进一步研究主要集中在如何消除在实现本节所描述的两种方案时遇到的问题。这些问题包括由于移动台运动造成的增益系数波动的影响,在这里是假设不变的;实现这些方法所需基站之间的通信;这些方法所需的同步执行;Foschini 和 Miljanic (1993)提出的改进分布式功率控制算法去除了通信的要求;改进方法只依赖本地测量,而且作者证明了这种方法的收敛性。Mitra 还证明了改进分布式功率控制算法非同步执行也能够收敛,这样,移动台之间就不需要假设同时执行算法迭代(Mitra,1994)①。

Foschini-Miljanic 的分布式功率控制的改进算法将式(4.29)用下面的等式替换

$$P_i^{(n)} = \gamma P_i^{(n-1)} / \gamma_i^{(n-1)} \tag{4.31}$$

用参数 γ 替换式(4.29)中的 $c^{(n-1)}$,它是一个"目标"SIR,是所有移动台都希望达到的。这个目标 SIR 必须由 $\gamma \leqslant \gamma^*$ 定义。所有移动台的迭代都在本地进行,不像利用 $c^{(n-1)}$ 那样需要移动台间的通信。通过重写出分贝形式的迭代表示,可以发现式(4.31)中有一个有趣的变化。注意,$P_{\mathrm{dB}} = 10\lg P$,立刻得到

$$P_i^{(n)}{}_{\mathrm{dB}} = P_i^{(n-1)}{}_{\mathrm{dB}} + \left(\gamma_{\mathrm{dB}} - \gamma_i^{(n-1)}{}_{\mathrm{dB}}\right) \tag{4.31a}$$

这个表达式的解释比较有趣。它是说把一个移动台测得的 dB 数和目标 SIR 的 dB 数之差加到以前发送的功率估计上,以获得新的估计。Blom 和 Gunnarson(1998)②证明了如果采用 dB 形式,文献中提出的大多数分布式功率控制方案都可写成式(4.31a)的形式。这些方法的统一形式可以通过在式(4.31a)中的差分项之前乘以一个适当选择的随迭代而变化的参数来获得。

本节开始就提到,这里讨论的方法适用于复用限制起作用的时分多址/频分多址系统。这里集中讨论的一维例子实际上假设的是 2 小区复用。在本节开始也提到这种方法的一种形式是作为第 6 章和第 8 章将要讨论到的 IS-95 码分多址系统的上行(移动台为发送端)功率控制方案。码分多址系统的目标是不论移动台位于小区内的什么位置,都确保每个移动台在某个基站测得的 SIR 值相同。在第 6 章中将会看到这个条件会带来最好的码分多址性能。为达到

---

① Mitra, D. 1994. "An asynchronous distributed algorithm for power control in cellular systems," in *Wireless and Mobile Communications*, ed. J. M. Holtzman, Boston, MA, Kluwer/Academic.

② Blom, J. and F. Gunnarson. 1998. "Power control in cellular systems, Linköping studies in science and technology," Thesis No. 706, Department of Electrical Engineering, Linköpings Universitet, Linköping, Sweden."

这个目标,通常采用两部分的功率控制模式(Viterbi,1995)[①]。一种是粗略的"开环"功率控制。在这种方法中,每个移动台根据其接收到的基站的下行功率进行自动增益控制,用所得的测量结果来调整其发送功率。但由于上行和下行传播条件可能有很大的不同,这只是一种粗略的调整。这里隐含的意思是式(4.12)中定义的上行和下行增益参数 G_{ij} 和 G_{ji} 是相同的。很明显,这只是一种粗略的近似。码分多址系统为了使基站所有的接收信号的 SIR 都有良好的质量,就必须结合一种"闭环"的功率控制方案。IS-95 所采用的用来控制其小区内每一个移动台发送功率的"闭环"功率控制策略是属于"bang-bang"型的:如果由基站估计的接收 SIR 低于一个确定的门限,基站会命令移动台将发送功率增加一个固定的 dB 值;如果由基站估计的接收 SIR 高于这个门限(Viterbi,1995),基站会命令移动台将发送功率减少一个固定的 dB 值。特别地,采用本节中的表达方式,有

$$P_i^{(n)}{}_{\mathrm{dB}} = P_i^{(n-1)}{}_{\mathrm{dB}} \pm \Delta \tag{4.32}$$

这里 Δ 是一个单位为 dB、固定的增量(例如,1 dB)。若估计的 $SIR_i < \gamma$,则取 + 号;若 $SIR_i > \gamma$,则取 − 号。但通过比较式(4.32)和式(4.31a),注意到两种功率控制步骤是很相似的。事实上,如果式(4.31a)的差分项量化为 1 比特,则两种形式等价。这样,码分多址和时分多址/频分多址的功率控制策略就密切相关了。最基本的不同在于测量和估计 SIR_i 的方法。注意,码分多址方法在估计每个接收的 SIR 后只需要从基站向每个移动台传输 1 比特的数据。

习题

4.1 参见图 4.1,假设我们给每个小区分配了 10 个信道。

(a) 系统总共有多少信道?

(b) 将文中提到的动态信道分配方法定向信道锁定的信道借用用于本系统:将信道从 1 到 10 做一张信道表。应用文中提到的信道排序,将本地呼叫从 1 往上按优先级分配,将借给相邻小区的信道从 10 往下按优先级分配。集中讨论一个小区。考虑正在进行通话数目各不相同的应用,这些通话包括本地的和借用相邻小区信道的,计算这种方法在不同应用下的情况。用这个例子说明信道锁定,然后说明定向锁定能帮助缓解这个问题。列举一些例子来说明文中描述的快速信道分配的特点。

(c) 用这种方法和其他方法进行比较,其他方法可以从文献中获取,也可以自行设计。

4.2 将上面的习题 4.1 用于图 4.2 中一维的情形。

(a) 如果每个小区中分配了 10 个信道,系统总共有多少信道?

(b) 说明在不同数目的呼叫进行的情况下,定向信道锁定的信道借用方法在这个系统中是如何实现的。

(c) 考虑小区 i 中有 10 个正在进行的呼叫。假设小区 $i-1$ 中有 7 个正在进行的呼叫。小区 i 中有一个用户试图发起呼叫。那个小区的基站从小区 $i-1$ 中借取了一个信道。使用定向信道锁定的信道借用方法,则现在有多少个信道锁定?是哪些信道?通过例子说明文中提到的信道借取不会传播。

---

① Viterbi, A. J. 1995. *CDMA, Principles of Spread Spectrum Communication*, Reading, MA, Addison-Wesley.

4.3 计算在系统有 $2m=2$ 个信道的情况下一维定向信道锁定的信道借用方法的阻塞概率。假设业务均匀，证明式(4.8)的结果。做出以厄兰为单位的业务强度 A 和阻塞概率的关系图，并和固定信道分配结果相比较。

4.4 考虑在 $2m=4$ 个信道的情况下，一维定向信道锁定的信道借用系统的情形。

(a) 证明总共有 55 个状态。

(b) 标出 25 个阻塞状态中的一些，解释它们为什么是阻塞状态。

(c) 假设业务均匀，将阻塞概率作为厄兰负载 A 的函数，计算并作图，与图 4.5 做比较。

4.5 参考图 4.7 和图 4.8。将功率控制的分布式平衡算法应用于每个例子。

(a) 找到每个例子中的 $\boldsymbol{R}$ 矩阵。特别地，证明式(4.26)为图 4.7 例中的 $\boldsymbol{R}$ 矩阵。

(b) 找到每个例子中的 $\boldsymbol{R}$ 的 6 个特征值，并证明它们与文中出现的数值一致。

(c) 证明各个例子中的 $\mathrm{Tr}(\boldsymbol{R})=M=\sum_i z_i$。

(d) 对每个例子都利用分布式平衡算法，并证明这种方法会收敛到文中提到的两个 SIR 值，18.7 dB 和 18.5 dB。对一些基站的 SIR 值和迭代次数 n 的函数关系画图，至少进行 20 次迭代；对结果进行解释。

4.6 这个问题与式(4.29)中定义的分布式功率控制算法的分析有关。

(a) 将 $\boldsymbol{A}$ 按文中定义，证明由式(4.29)可将这种算法的矩阵向量形式写成式(4.30)的形式。

(b) 用 $\boldsymbol{A}$ 矩阵的特征值得到分布式平衡算法中定义的 $\boldsymbol{R}$ 矩阵的特征值。

(c) 证明分布式功率控制算法收敛到与分布式平衡算法相同的 SIR。说明这种方法为什么会比分布式平衡算法收敛快很多。

4.7 仍考虑分布式功率控制算法的习题 4.5。找到图 4.7 和图 4.8 中的矩阵 $\boldsymbol{A}$。计算每个的特征值，与分布式平衡算法的特征值进行比较。将每个例子按算法至少进行 20 次迭代并作图，再与分布式平衡算法相比较。

4.8 自行选择一个至少由 13 个小区组成的一维线性蜂窝网络，将分布式平衡算法和分布式功率控制算法做比较。

(a) 对两小区复用和 3 小区复用的情况进行迭代分析，做出一些基站的 SIR 和迭代参数的曲线。尝试选择最差情况下的移动台位置。

(b) 将分布式功率控制算法收敛到 SIR 和分布式平衡算法中矩阵 $\boldsymbol{R}$ 的最大特征值相比较。

(c) 得到每个移动终端的相对功率，并说明结果。

4.9 考虑式(4.31)中定义的 Foschini-Miljanic 算法。

(a) 与式(4.29)中定义的分布式功率控制算法相比较，证明这两种方法收敛到相同的 SIR。

(b) 说明在执行这种方法时为什么不需要移动台之间的通信。

(c) 将这种方法用于图 4.7 中的 11 个小区的线性系统，并且与文中的结果相比较。

4.10 这个习题的目的是提供式(4.32)中固定增量功率控制的一个应用实例。从一个自行选择的小区网络开始(简单的可以选择正文例子中采用的一维线性网络)。让每个可能干扰的移动台选择一个初始功率。每个移动台的基站估计其 SIR，然后按文中讨论的那样，命令移动台减少或增加一个固定的增量。将这个步骤迭代进行直到收敛达到。按不同增量值使用该方法，并讨论结果。

第5章 调制技术

在这本书的前一部分，着重讲述了移动蜂窝系统的基本系统概念，讨论了传播环境及其建模，以及按地理位置划分小区并通过信道空间复用来增加系统容量的方法。在前一章中，又讲述了如何用动态信道分配和功率控制的方法来提高系统的性能。所有这些都暗含着一个假设，即用户是在分配的信道上进行通信。根据所研究的系统不同，信道可以是频带、频带的时隙或者一种特定的编码(一个给定的二进制符号的序列)。在这些情况中，通信是在特定频率的载波上用不同的调制技术来传送想要的信息的。所传输的可能是数据、语音或者其他种类的在载波上可携带信息的信号。在这一章中，我们将描述各种已提出或已使用的、能最有效利用信道的调制技术。

通信系统中广泛使用载波频率的调制或变换来使信息在需要的载波频率上传输或进行特定的操作。在无线移动系统中，调制过的无线信号即可以按情况在上行链路(从移动台到基站)或下行链路(从基站到移动台)中传输。前一章所讨论的传输功率就是已调信号的功率。第一代蜂窝系统，比如第1章中提到的AMPS系统，使用模拟FM作为调制方案。第二代系统，如数字AMPS或D-AMPS(也称为IS-54/136)、IS-95或者GSM，使用的是数字调制方案，这些系统将在第6章中介绍，并在第8章中描述其具体细节。将在第10章中讨论的第三代系统，以及将在第12章中讨论的个人通信网络，也使用的是数字调制。用第二代系统作为例子。在第1章中提到过，美国移动蜂窝系统的频带分配在上行链路是824 MHz到849 MHz，在下行链路是869 MHz到894 MHz。AMPS和D-AMPS都把各方向上的25 MHz可用频带分成832个30 kHz的信道。我们还看到，D-AMPS在每个30 kHz的信道中又通过使用时隙使系统容量增加了3到6倍。D-AMPS选择的调制技术是$\pi/4$-DQPSK。美国IS-95使用了相同的上行和下行链路频带，但它使用码分信道，每个信道占用1.25 MHz带宽，比前两者宽得多。它在下行传输中使用QPSK调制，在上行链路传输中使用OQPSK调制，这两种调制技术都会在本章中讲到。

在欧洲，分配给GSM使用的带宽在上行方向是890 MHz到915 MHz，下行方向是935 MHz到960 MHz。可用的25 MHz的信道进一步分成124个200 kHz的频道。每个频道上有8个用户通过使用分配的时隙轮流接人。使用的调制技术为0.3 GMSK，这在本章后面也会讲到。

蜂窝通信的附加无线通信信道在美国可以使用1.85 GHz到1.99 GHz频带；称为PCS频带，在欧洲则是1.71 GHz到1.9 GHz。本章中为了简化示例，主要关注800 MHz到900 MHz带宽的调制技术。

本章中的讨论将主要集中在上面提到的数字调制技术。先对数字调制技术做一些介绍，然后再详细描述DQPSK，GMSK，QPSK和OQPSK等具体调制方法。我们也可看到，在选择调制技术时需要考虑多个标准。这些标准包括带宽利用率、功率效率、费用和复杂度以及在衰落信道中的性能。特别是使用电池供电的移动台或用户终端必须价格适当、体积小并且节能。在衰落环境中则意味着倾向于使用非线性放大器的恒定包络调制技术，这也是第一代模拟系统最初选择FM调制的原因。

本章中将首先描述最简单的基本的数字调制技术。包括开关键控(OOK)或幅度键控

(ASK)、相移键控(PSK)以及频移键控(FSK)。然后,将简要了解一下正交幅度调制(QAM),这种调制主要用在有线调制解调器中,以提高在带宽有限信道中的传输比特率,比如从家庭或办公室到电话交换机或中心办公室的接入线。正交相移键控(QPSK)是 QAM 的特殊情况。然后,会继续讨论增强的数字调制技术如 OQPSK 或 MSK 以及 GMSK,它们在无线蜂窝系统的受限带宽以及衰落环境中工作相对良好。

最后将介绍一下正交频分复用(OFDM),这种至少可回溯到 20 世纪 60 年代的方案目前在高级的蜂窝系统中使用。它已经用于高速无线网和通过电话铜线接入家庭的高速 DSL 调制解调器中。我们将在第 12 章中介绍它作为 IEEE 802.11g 和其他标准的一部分在无线局域网中的应用。

5.1 数字调制技术介绍

考虑一个未经调制的连续正弦载波 $A\cos\omega_0 t$,其频率为 f_0 Hz 或 $\omega_0 = 2\pi f_0$ rad/s。该载波的幅度、频率和相位可能分别被携带某种信息的信号所调制,前两者就产生了人们所熟悉的在模拟无线广播系统中常用的 AM 和 FM 高频信号。前面提到的最简单的数字调制被分别称为开关键控或幅度键控传输(OOK 或 ASK)、频移键控传输(FSK)和相移键控传输(PSK)。在数字调制中,连续二进制数字(比特)0 或 1,用来改变或者调制未调制载波各自的幅度、频率和相位。这些携带了信息的二进制数字同样可看成是组合基带信号序列。在最简单的 OOK 情况中,基带信号序列中的二进制 0 会将载波截断,而二进制 1 出现的地方则有 $A\cos\omega_0 t$ 传输。设二进制序列的传输比特率为 R bps,则每个二进制符号或比特持续 $1/R$ s。图 5.1 中表示了对应于 3 个连续比特 1,0,1 的 OOK 信号。注意 OOK 信号不适合数字无线系统的传输,因为其传输信号的幅度有变化。前面说过,恒定包络的信号更适合衰落环境,因为它使信号幅度产生不可知的随机变化。在这里通过讨论 OOK 传输来使大家对数字通信有一个全面的认识。

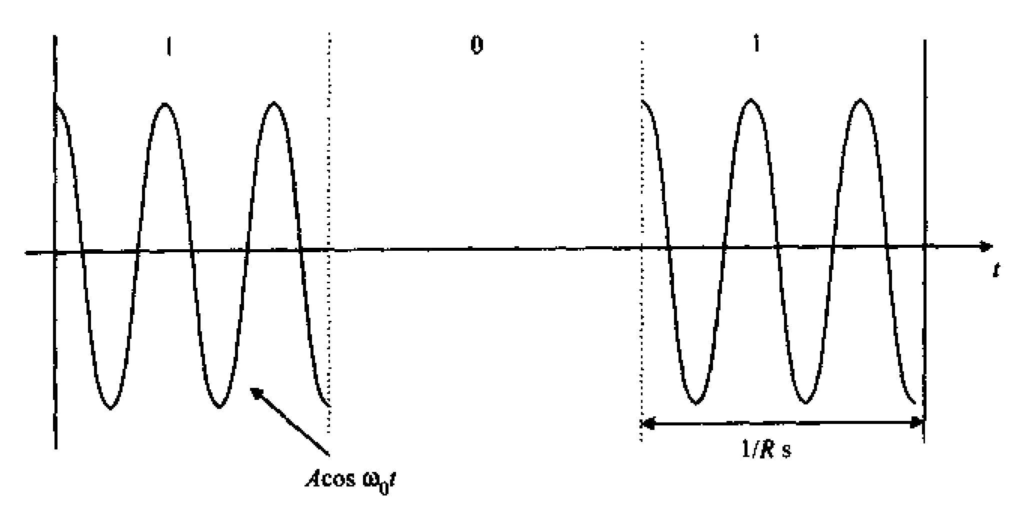


图 5.1 二进制信号 1,0,1 的对应 OOK 调制信号

事实上,图 5.1 显示的 OOK 信号突然消失和再次出现的情况在现实中不可能发生,因为这么陡峭的转换需要无限的带宽。实际操作中,传输信号一般需经过成形或预滤波才能在 0 和 1 之间进行更加平缓的过渡。对应于基带序列 1 的传输信号可以写成 $Ah(t)\cos\omega_0 t$。低频时间函数 $h(t)$代表一个信号成形函数,设计这个函数使得允许以指定的系统带宽传输 OOK 序列,而没有失真或只有很小的失真。关于信号成形会在下一节中讨论。RF 载波传输带宽,

也就是载波调制的传输信号的频谱容易给出为 $2B$，B 是基带带宽，即传输二进制的 0，1 序列所需要的带宽(Schwartz，1990)[1]。在下一节我们会看到，基带带宽的单位是 Hz，随着这里的成形类型的不同而在 $R/2$ Hz 和 R Hz 之间变化，因此传输信号带宽在 R Hz 和 $2R$ Hz 之间变化。例如，如果在无线信道中传输一个 10 Kbps 的数字信号，基带带宽根据使用的成形类型的不同在 5 kHz 到 10 kHz 之间变化，而无线信道的传输带宽则是从 10 kHz 到 20 kHz。由于是 OOK 传输，即使有信号成形，也还是会有幅度变化的，因此对于移动通信不满足要求，需要其他类型的数字调制来提供所需数字信号的调制。实际中，大部分是使用各种相移键控即 PSK 来实现的。但是，先来讨论频移键控 FSK，这是 FM 的二进制形式。对 FSK 的简要讨论会对后面讨论更多形式的数字通信调制形式有帮助。

FSK 中，载波频率在两个不同的频率 $f_1 = f_0 - \Delta f$ 和 $f_2 = f_0 + \Delta f$ 之间转换，取决于传输的是 0 还是 1。术语 Δf 被称做关于平均载波频率 f_0 的频率偏差，若用角频率表示，则 $\omega_1 = \omega_0 - \Delta\omega$ 和 $\omega_2 = \omega_0 + \Delta\omega$，每个角频率是对应的用 Hz 表示的频率的 2π 倍。具体地，在一个长 $1/R$ s 的二进制间隔中，设 $A\cos\omega_1 t$ 对应于传输的 1，而 $A\cos\omega_2 t$ 对应于传输的 0(严格来讲，每个载波都要乘以成形项 $h(t)$，但为了简单起见将这些项忽略)。图 5.2 中有一个例子，传输信号带宽大约是 $2\Delta f + 2B$，其中基带带宽 B 根据使用信号成形方法的不同而在 $R/2$ 和 R 之间变化(Schwartz，1990)。FSK 传输带宽比 OOK 传输需要的带宽大很多。对于模拟 FM 调制也有相同的结果。在图 5.3 中表示了一个简单的 FSK 频谱，其中指出了载波频率和传输带宽。

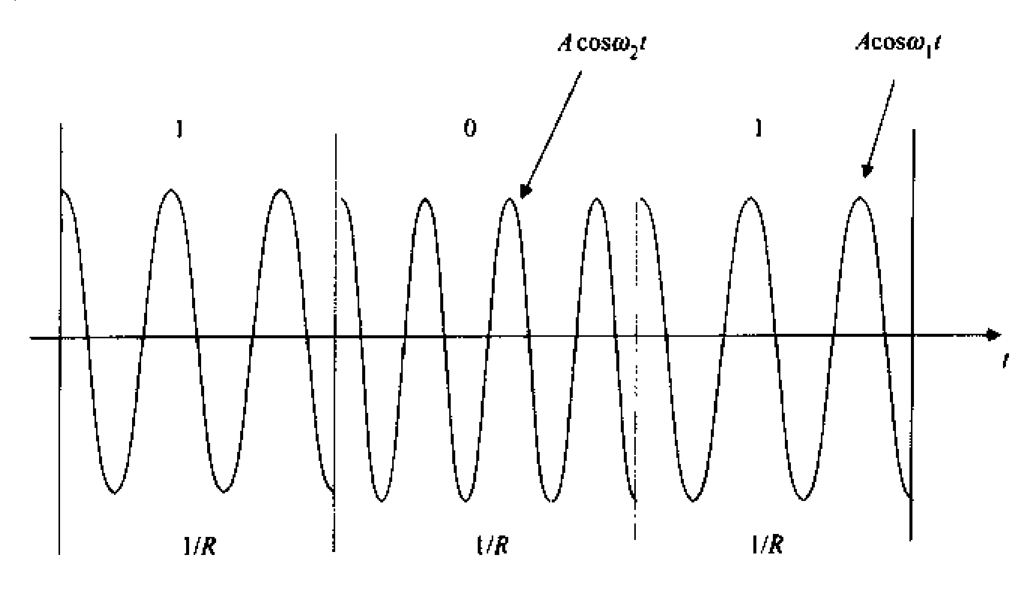


图 5.2 FSK 调制信号

下面来看相移键控 PSK，这是对承载着信息的二进制数字序列进行载波调制的 3 种基本方法的最后一种。在这种情况中，举个例子，当基带二进制序列中有 1 出现时载波项是 $A\cos\omega_0 t$，当有 0 出现时载波项是 $-A\cos\omega_0 t = A\cos(\omega_0 t + \pi)$(为了讨论简便，再一次忽略成形因素)。图 5.4 表示了一个例子。PSK 传输是噪声存在情况下最好的调制方案(Schwartz，1990)。然而，为了保证准确无误地检测出传输的二进制序列，在接收端需要有相位参考。注意，FSK 和 PSK 都是恒定包络的传输，这是伴随着随机幅度变化的信号在衰落环境下需要的特

[1] Schwartz, M. 1990. *Information Transmission, Modulation, and Noise*, 4th edn, New York, McGraw-Hill.

征。从图5.4中可以看出信号从1到0变换时会发生陡峭的相位变化(这种效果在图5.4没有信号成形的情况下更加明显)。相位上的陡峭变化(图5.4中的 π 或 $180°$)在传输过程中会被一定程度地修改和平滑化,因为陡峭的变化意味着信号频谱上无限的带宽,这在实际中是不可能的。然而,相位的变化一定会导致不需要的信号传输带宽的增加和相应的由于有限的传输带宽而导致的传输信号幅度的变化。因此,必须采取一定的措施来减小这些大的陡峭的相位变化。另外,我们也会讨论这样的调制技术的修正措施。在这之前,在下一节中先定量讨论信号成形的问题。这个问题很重要,因为在后面将要讨论的适用于无线移动通信系统的增强调制技术方面,信号成形的概念发挥着关键的作用。

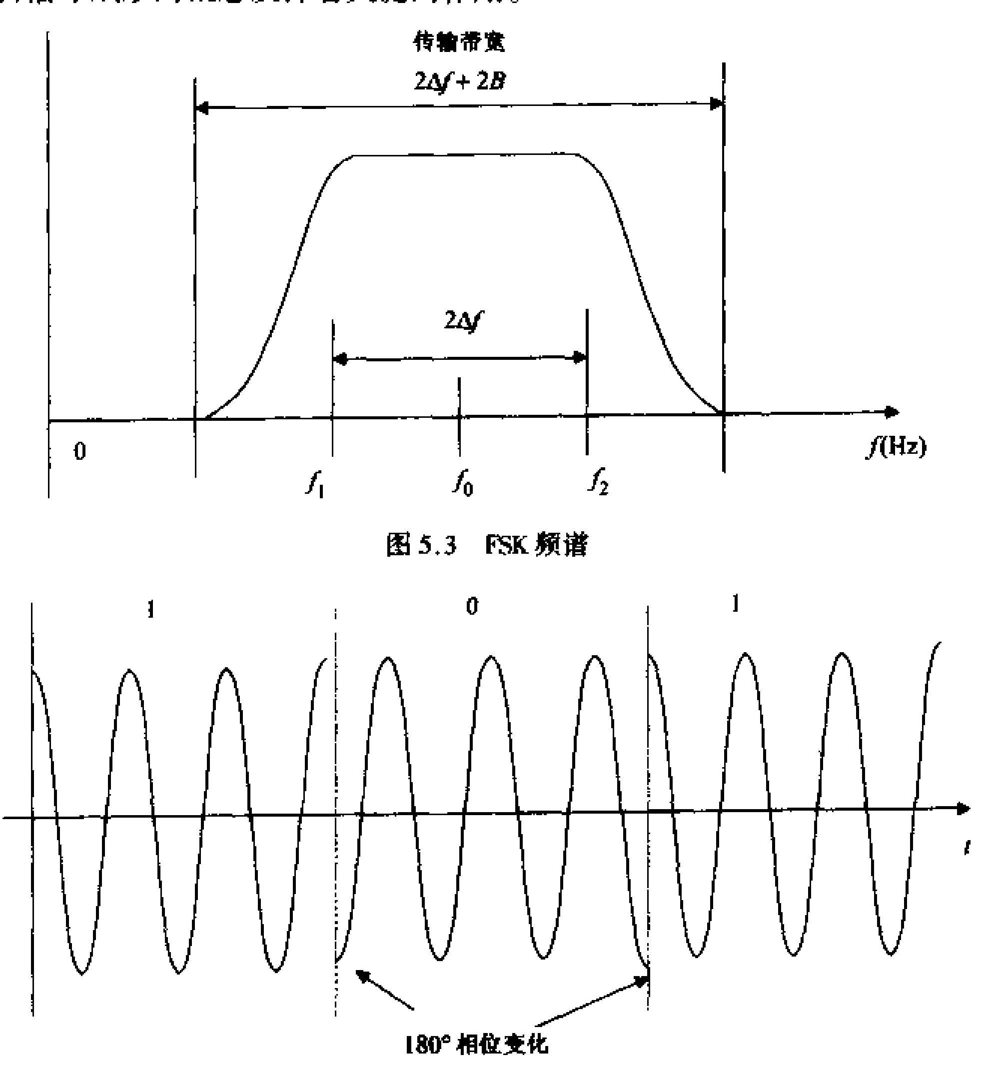


图5.3 FSK频谱

图5.4 PSK信号

5.2 信号成形

在前一节中讲到,在数字通信方案中必须采用信号成形来保证调制的传输信号的带宽在指定系统带宽的范围内。在这一节我们将继续讨论成形,主要关注在数字通信系统设计中使用了多年的一系列成形函数。其他类型的成形函数会在下一节讨论移动通信系统特定的调制方案时遇到。这一节的讨论会对信号成形有大致的理解。

回顾前面一个高频频率为 f_0 的无线载波信号,可以写成 $h(t)\cos\omega_0 t$ 的形式。时变函数

$h(t)$即为成形函数。为了保持调制信号的平滑，尤其是在从1到0的转换期间保持平滑，显然，函数$h(t)$必须有图5.5所表示的形式。函数在二进制间隔$1/R$的中间有最大值，然后在最大值的两边平缓地下降。这样，不管是0还是1都可以通过中心值抽样来再生原来的二进制符号。可以看到，如果这样的一个函数乘以图5.1中的OOK正弦波，则可以明显地减小从1到0转换时幅度的陡峭变化。先不考虑对图5.1、图5.2和图5.4中高频调制信号以及将它们乘以函数$h(t)$效果的分析，只考虑基带二进制信号0,1序列的成形情况会更加简单。乘以$\cos\omega_0 t$会使成形后的频谱移动到以载波频率f_0为中心的频谱上。为了概念上的简单，成形可以在基带实行，也可以直接在载波频率上实行，其结果是一样的。

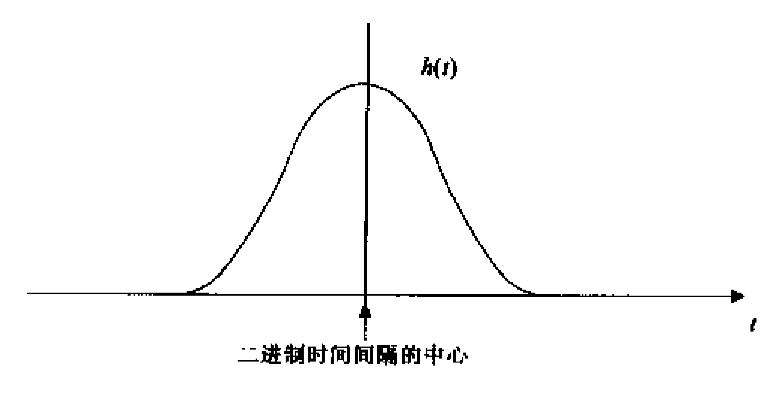


图5.5 典型成形函数

考虑一个基带0,1序列，对应的比特率是R bps，每个比特乘以$h(t)$，如图5.6所示。比特1用$h(t)$函数的出现来表示，而比特0则表示时间序列上的空白。从傅里叶分析可以得到，如果一个时间函数变窄，则在频域上其频带会相应地增加。而若函数在时域上加宽，则频域上带宽就会减小。脉冲的宽度和它的带宽彼此互为反相关关系(Schwartz,1990)。因此，如果图5.6中的$h(t)$宽度减小，则基带信号序列的带宽就会增加；而$h(t)$的宽度增加，则相应的带宽就会减小。然而，随着$h(t)$宽度的增加，脉冲开始扩展到邻近的二进制间隔。这就会引起符号间的干扰，如图5.6所示。最初没有问题，因为总是能够通过在脉冲的中心抽样来决定是0还是1。但是，当最后符号间干扰大到影响了二进制信号在其时间间隔中心的值时，把1混淆成0就可能发生。当存在噪声干扰传输信号时，这种情况更加明显。这样就存在着一个在符号间干扰和传输脉冲序列所需的带宽之间的权衡折中。同样的结论也适用于乘以$\cos\omega_0 t$并移动到载波频率f_0的调制二进制序列。在下一节中我们会看到这个折中也出现在数字移动通信系统使用的特定的成形函数中。

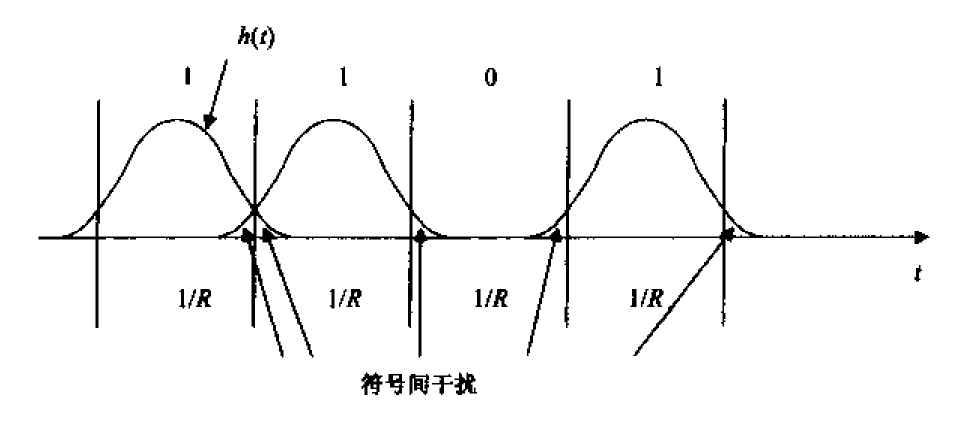


图5.6 二进制序列

为了更加具体地阐明带宽和符号间干扰这一权衡关系，现在讨论在许多数字通信系统中使用的一种特定类型的成形。在当前的第二代数字移动通信系统中使用的其他类型会在下一节中讲述。这种成形称为正弦滚降成形。为了介绍这种成形，先来看一种特殊情况，升余弦成形。使用升余弦这个词是因为函数 $h(t)$ 的傅里叶变换或频谱 $H(\omega)$ 是升余弦形状，如下式表示：

$$H(\omega)=\frac{1}{2}\left(1+\cos\frac{\pi\omega}{2\omega_c}\right) \quad |\omega|\leqslant 2\omega_c \tag{5.1}$$
$$=0 \quad 其他$$

这个函数的图形在图 5.7 中大致给出。这种成形类型和正弦滚降成形都是 Nyquist 成形函数中的特殊例子，这类函数以 AT&T 伟大的研究者 H. Nyquist 命名，他于 20 世纪 20 年代在对电报的研究中第 1 次描述了这些成形函数（Nyquist，1928[①]；Schwartz，1990）。注意到式（5.1）的升余弦频谱和图 5.7中的带宽是 $2\omega_c$，单位是 rad/s，ω_c 是一个可调整的参数。用 Hz 单位表示的带宽 B 可以得到 $B=\omega_c/\pi$，这在图 5.7 中已经指出。用 OOK 或 PSK 传输的高频或 RF 带宽是这个值的两倍。定义变量 $\Delta\omega\equiv\omega-\omega_c$，$\Delta\omega$ 用来表示与变量 ω 的偏差，如图 5.8 所示。我们可以得到式（5.1）的另一种写法，在式（5.2）中给出，它更直接地接近下面讨论的正弦滚降成形的一般情况。

$$H(\omega)=\frac{1}{2}\left(1-\sin\frac{\pi}{2}\frac{\Delta\omega}{\omega_c}\right) \quad -\omega_c\leqslant\Delta\omega\leqslant\omega_c \tag{5.2}$$

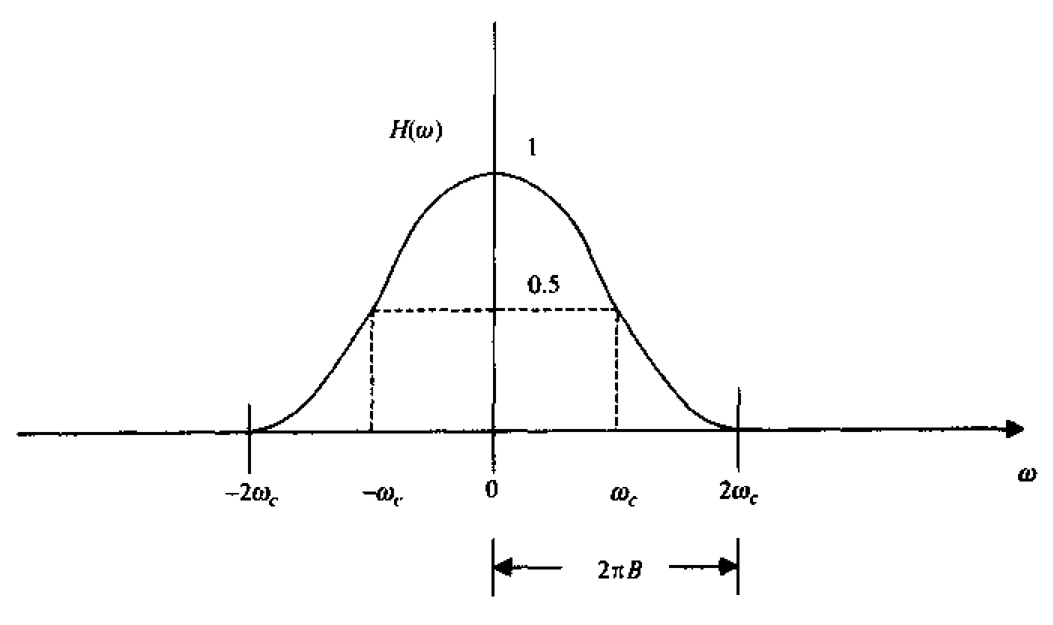


图 5.7 升余弦频谱

取式（5.1）或式（5.2）的反变换，可得到脉冲函数或成形函数 $h(t)$，在时域可写成

$$h(t)=\frac{\omega_c}{\pi}\frac{\sin\omega_c t}{\omega_c t}\frac{\cos\omega_c t}{1-(2\omega_c t/\pi)^2} \tag{5.3}$$

这个函数的大致形状如图 5.9 所示。函数具有以 π/ω_c s 的间隔穿过 0 的特性。以这个间隔分开并以其他数列的 0 点为中心的二进制数列不会互相干扰。图 5.9 中，在主脉冲两边用以

① Nyquist, H. 1928. "Certain topics in telegraph transmission theory," *Transactions of the AIEE*, 47 (April). 617 - 644.

π/ω_c 为间隔的两条虚线表示的脉冲表示了这个原理。在实际中，一些脉冲抖动总会出现，从而导致符号间的干扰。但从式(5.3)中可以看到这些二进制脉冲的尾部在 t 足够大的情况下会衰落 $1/t^3$，因此，任何符号间干扰都可以降低到一个可以容忍的水平。选择二进制间隔 $1/R$ 作为$h(t)$零点之间的间距，则有 $1/R=\pi/\omega_c$。二进制传输速率 R 所需基带的带宽用 Hz 表示，即$B=\omega_c/\pi=R$。举个例子，以 14.4 Kbps 传输的数据流，如果使用升余弦成形，则所需基带的带宽是 14.4 kHz。相应地，以载波频率为中心的高频带宽对于 OOK 或 PSK 传输则是 28.8 kHz。加倍传输速率会加倍这两种数字高频传输所需的传输带宽。

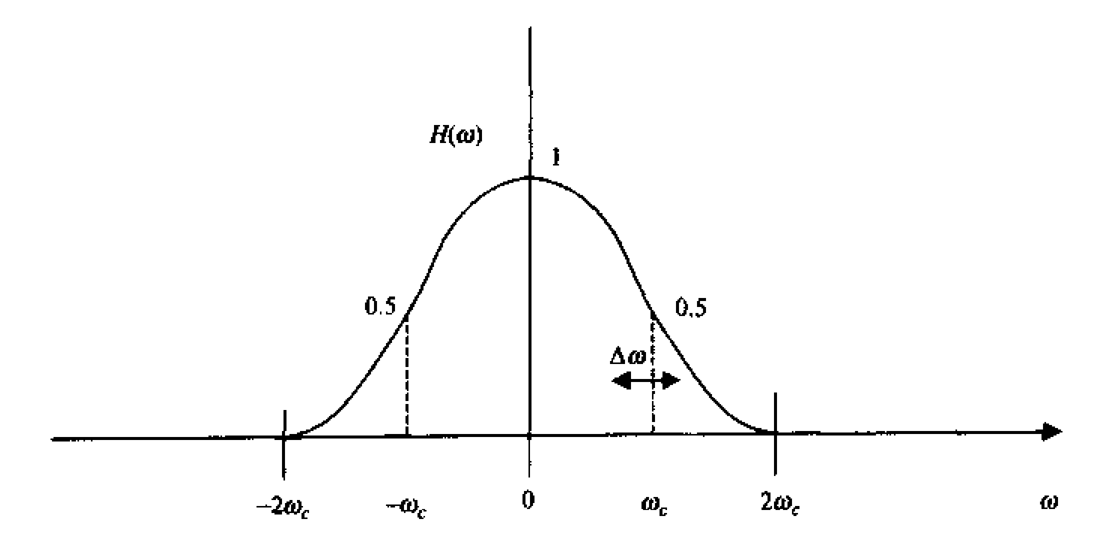


图 5.8 替代变量的图示

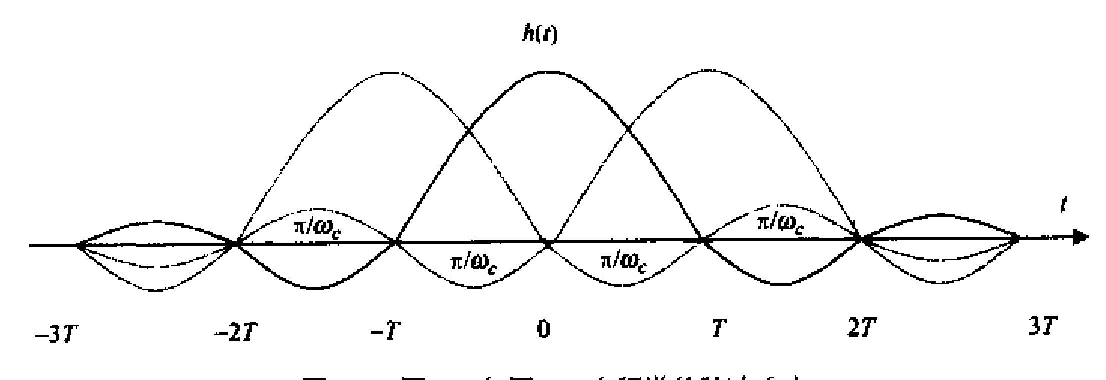


图 5.9 图 5.7 与图 5.8 中频谱的脉冲响应

现在来看正弦滚降这种更一般的情况。这种方法可以控制符号间干扰，并保持传输所需的带宽尽量低。上面提到过，升余弦成形是一种特殊的形式。正弦滚降成形则是通过下面的频谱 $H(\omega)$成形的：

$$
\begin{aligned}
H(\omega) &= \frac{1}{2}\left(1-\sin\frac{\pi}{2}\frac{\Delta\omega}{\omega_x}\right) && |\Delta\omega|\leqslant\omega_x \\
&= 1 && -(2\omega_c-\omega_x)<\Delta\omega<-\omega_x \\
&= 0 && |\Delta\omega|>\omega_x
\end{aligned}
\tag{5.4}
$$

频谱形状如图 5.10 所示。注意到其中引入了一个参数 $\omega_x\leqslant\omega_c$ 替代了式(5.2)的升余弦频谱中出现的参数 ω_c，因此使升余弦成形频谱得到了推广。在式(5.4)中令 $\omega_x=\omega_c$ 就可以得到式(5.2)。再将式(5.4)中的变量定义成$\Delta\omega=\omega-\omega_c$，相应的脉冲波形 $h(t)$，即 $H(\omega)$的傅里

叶反变换可以表示成

$$h(t)=\frac{\omega_c}{\pi}\frac{\sin\omega_c t}{\omega_c t}\frac{\cos\omega_x t}{1-(2\omega_x t/\pi)^2} \tag{5.5}$$

这个公式与式(5.3)中升余弦脉冲响应具有相同的形式，只是在表达式的第 2 部分参数 $\omega_x\leqslant\omega_c$ 代替了参数 ω_c。图 5.10 中弧度表示的带宽是 $2\pi B=\omega_c+\omega_x$。若以 Hz 为单位，则 $B=f_c+f_x$，其中两个频率参数 f_c 和 f_x 是将相应的弧度频率项除以 2π 得到的。带宽 B 比在升余弦中需要的带宽小，只有在 f_x 接近 f_c 时才比较接近。式(5.5)中的 $\sin\omega_c t/\omega_c t$ 项表明连续的二进制数列仍可能以 $1/R=\pi/\omega_c$ 为间隔，但符号间的干扰为 0。然而，由于 $f_x=\omega_x/2\pi$ 减小了，从而带宽 B 减小。我们必须更加严格地控制时域抖动，因为 $h(t)$ 第 2 项分母的 $1/t^2$ 衰减比升余弦的情况变慢了，$(\omega_x t)^2$ 项比式(5.3)中相应的项小，因此就存在一个带宽和时域抖动控制的折中。f_x 增加，逼近 f_c，即其在升余弦情况中的值，则由时域抖动引起的符号间干扰的可能性就会降低，但是带宽仍会增加。图 5.9 是这种更加一般的成形的表示形式。

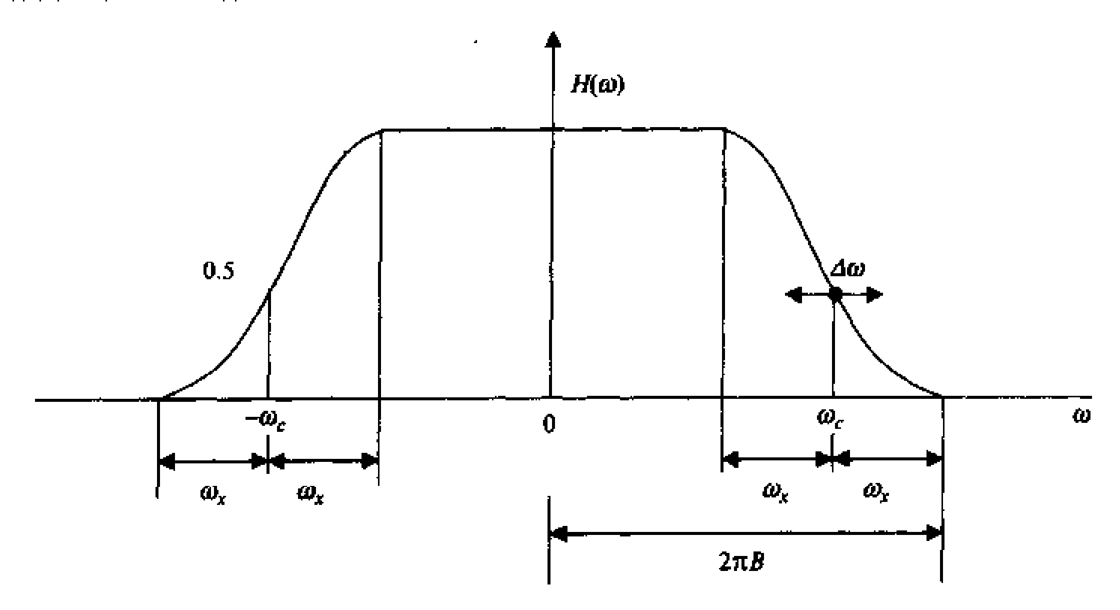


图 5.10 正弦滚降成形频谱

传输这些脉冲需要的传输带宽 $B=f_c+f_x$ 一般写成包含 f_x/t_c 比例的形式，这个比例定义成滚降因子 r。从图 5.10 中，我们可以看出这个比例决定了成形频谱从最大值 1 到 0 的下降速度。当 f_x 或相应的 r 减小趋近于 0 时，频谱下降得更快，带宽 B 减小，但是，时域抖动带来的符号间干扰的可能性却增大了。回顾一下，传输比特率是 R，从图 5.7 中可得到 $R=\omega_c/\pi=2f_c$，因为 $\omega_c=2\pi f_c$。这样，我们可以把带宽 B 写成下面的形式：

$$B=f_c+f_x=(R/2)(1+f_x/f_c)=(R/2)(1+r) \tag{5.6}$$

这里，滚降因子 $r\equiv f_x/f_c\leqslant1$。对于 OOK 或 PSK 传输的无线带宽，有 $2B=R(1+r)$。在这些情况中，当 $r=1$ 时带宽的值从刚大于 R 变成了 $2R$。令传输带宽为 B_T，则有

$$B_T=R(1+r) \tag{5.6a}$$

在随后对传输带宽的讨论中，经常去掉下标，用字母 B 简单地来表示传输带宽。标号的变化不会引起混淆，在讨论中会注明是基带还是传输带宽。

举个例子,假设进行 14.4 Kbps 的传输。如果使用正弦滚降成形和 PSK 调制的话,若滚降因子是 0.5,则传输带宽是21.6 kHz;若滚降因子是 1,则传输带宽是 28.8 kHz。如果传输率是 28.8 Kbps,那么带宽加倍。相反地,给定可用带宽,则得到对于给定滚降因子的最大二进制传输速率。第二代蜂窝无线系统 D-AMPS(IS-136)使用正弦滚降成形,滚降因子是 $r=0.35$。系统将全部的带宽分成若干30 kHz 的用户带宽。如果使用 PSK 的调制方式,则最大的二进制传输率是 $R=30\ 000/1.35=22.2$ Kbps。实际使用的调制方式将会在下一节中描述,称为 DQPSK。实际的传输率是 48.6 Kbps。日本 PDC 系统也使用正弦滚降成形,滚降因子是 $r=0.5$。每个用户的可用带宽是 25 kHz。对于 PSK 调制,二进制传输率是 16.7 Kbps。这个系统实际也使用 DQPSK 调制,实际的二进制传输率是 42 Kbps。在下一节中可以看到,更复杂的调制方式是怎样允许更高的传输速率的。

5.3 蜂窝无线系统中的调制

从 5.1 节中知道,关于数字调制方案,在处理带宽有限的蜂窝无线信道的传输问题上会出现一些具体的问题,可能会带来由衰落引起的随机幅度变化,因此必须选择一种具有恒定包络的调制方案。在 5.2 节中介绍的 PSK 和 FSK 具有这个特性,但是在从一个比特值变化到另一个比特值时发生的瞬间相位至 π(180°)的变换,会引起传输中由于带宽受限带来的幅度变化。另外,在 5.2 节中,已知使用 PSK 调制,在带宽为 B 的受限信道上的比特率根据 B 而严格受限,并与 B 直接成比例。因此,要找一种更加复杂的调制方案,能够产生在比特变化时具有有限幅度和相位变化的传输信号,同时对于给定带宽允许更高的比特率。这种类型的调制方式,本节中给出了它们的一些例子,它们是基于并利用了相位正交传输的。

在本节的开始讨论了一种称为正交 PSK 或 QPSK 的技术,这种技术允许系统对于给定的带宽以原来比特率的两倍速率进行传输。考虑同时传输两个相同频率 f_0 的载波,一个载波与另一个正交:$\cos\omega_0 t$ 和 $\sin\omega_0 t$。简单起见,令它们都具有两个可能的幅度值 ± 1。每个载波持续时间长度相同为 T s,与 PSK 信号相对应。令 $a_i=\pm 1$ 为对应于第 i 个信号的余弦(同相)载波的值;令 $b_i=\pm 1$ 为对应于同一个信号的正弦(正交)载波的值。两个载波经成形函数 $h(t)$ 成形。第 i 个信号 $s_i(t)$,持续长度 T,可由下面得出

$$s_i(t)=a_i h(t)\cos\omega_0 t+b_i h(t)\sin\omega_0 t \quad 0\leqslant t\leqslant T \tag{5.7}$$

假设这里的成形函数 $h(t)$ 以间隔 T 的中点为中心。这个公式由三角函数也可以写成下面的幅度/相位形式:

$$s_i(t)=r_i h(t)\cos(\omega_0 t+\theta_i) \quad 0\leqslant t\leqslant T \tag{5.7a}$$

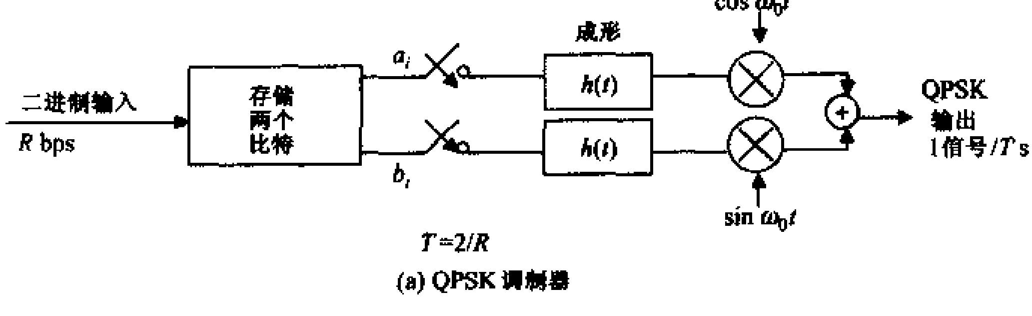
幅度-相位组(r_i,θ_i)与相应的(a_i,b_i)值相关联。因此,根据(a_i,b_i)的值,可以得到 4 个可能的信号,每个信号的传输需要相同的带宽。a_i 和 b_i 的实际值是如何确定的呢? 传输是如何进行的呢? 令比特率为 R bps。每 T s 连续的比特对被存储,产生 4 个不同的 2 比特序列。可以得到 2 比特间隔为 $T=2/R$。这 4 个序列中每一个被映射到 4 个载波信号中,从而特定的持续 T s 的信号被传输。这些信号可能的映射在表 5.1 中给出。合成的 QPSK 调制器的模块图如图 5.11 所示。

如果使用正弦滚降成形,如在式(5.4)或式(5.5)中,基带带宽 $B=(1/T)(1+r)$,用 $1/T$ 代替

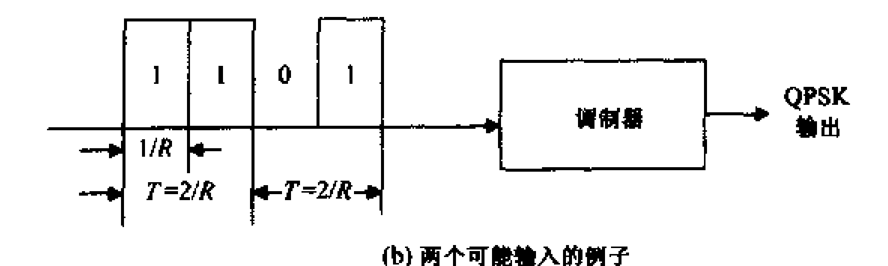
前面的 $R/2$,则传输一个特定比特序列所需的带宽减小了一半。相反地,给定带宽 B,可以得到两倍的传输比特率。PSK 情况中的传输带宽或者 RF 带宽是基带带宽的两倍,为 $(2/T)(1+r)$。解调即在接收端重构原始的二进制序列,在认识到式(5.7)中的 QPSK 信号的正弦项和余弦项互相正交之后,就可以很容易地进行了。将接收到的 QPSK 信号分别进行乘以余弦项和正弦项的操作,每个信号要与余弦和正弦严格同相,然后在 T s 的间隔内进行积分(将两个乘积项低通滤波),就可以分别提取出各项(Schwartz, 1990)。这个算法需要了解传输信号严格的相位,这与先前在 PSK 传输中遇到的问题相同。一个可替代的操作将会在本节后面简要地描述,这是利用连续信号作为后面信号的相位参考。

表 5.1 二进制输入序列→QPSK(a_i, b_i)

相连的二进制数对	a_i	b_i
00	-1	-1
01	-1	+1
10	+1	-1
11	+1	+1



(a) QPSK 调制器



(b) 两个可能输入的例子

图 5.11 QPSK 信号的生成

正交调制的操作可以通过一个二维图表 5.12(a)从几何上示意。横轴表示式(5.7)的同相余弦项系数 a_i;纵轴表示正交正弦项系数 b_i。将(a_i, b_i)的 4 个可能值放在一起就得到了图 5.12(a)中所示的 4 个点。这些点与式(5.7a)中(r_i, θ_i)的 4 个值直接对应。这些点在下面的讨论中称为 QPSK 信号的信号星座图。每个点可以写成复数形式,或者向量 $\boldsymbol{a}_i$,实部为 a_i,虚部为 b_i。分析一下便知道 QPSK 信号的另一个等同的表示可以通过旋转图 5.12(a)中的信号星座点 45°或 $\pi/4$ 来得到,从而产生图 5.12(b)的星座图。这个星座图对应于传输 4 个信号 $\pm\cos\omega_0 t$, $\pm\sin\omega_0 t$,取决于 $\boldsymbol{a}_i=(a_i, b_i)$的值。这代表了 4 个可能值的另一种映射方法,其例子在表 5.2 中给出。注意,尽管这个信号星座图中的 4 个点位于同一个圆周上,但这些点之间的转换及其信号的传输,也会在转换时间附近产生一些信号幅度的变化。

T s 的间隔对应于存储两个比特的时间，也是每个 QPSK 信号(4 种可能中的一种)传输的时间，这个时间称为波特或符号间隔，它与二进制间隔 $1/R$ 不同。它指的是一定数目的波特，而不是 bps，波特数决定了带宽。在二进制的特殊情况中，传输波特数和比特率是一样的。但是对于 QPSK 来说，它们之间差一倍。使用符号或波特率 $1/T$ 来代替 R 则必须修改式(5.6)和式(5.6a)。这个变化在前面已经说过。比如，假设用 QPSK 传输 28.8 Kbps，对应于每秒传输 14.4 k 波特或 14.4 k 符号/s。如果使用正弦滚降成形，则基带带宽是 $7.2(1+r)$kHz。传输无线 QPSK 信号所需的带宽是基带带宽的两倍，为 $14\,400(1+r)$Hz，其中 r 是滚降因子，与前面一样。

表 5.2　π/4 DQPSK

(a_i, b_i)	$\Delta\theta(n)=\theta(n+1)-\theta(n)$
-1, -1	$-3\pi/4$
-1, +1	$-\pi/4$
+1, -1	$\pi/4$
+1, +1	$3\pi/4$

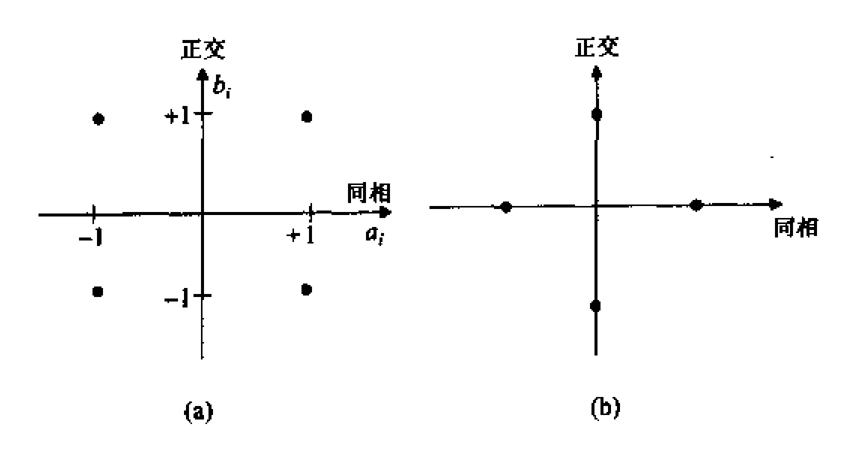


图 5.12　QPSK 信号星座图

现在把存储两个比特来减小所需带宽这个思路推广一下。举个例子，设 3 个连续比特被存储，结果是 $2^3=8$ 产生 3 比特序列。对应于这些序列，可以传输 8 种不同的信号，它们具有相同的幅度但是相位相差 π/4。这 8 个信号也可以相位和幅度都不同。前面的调制形式称为 8-PSK 调制。这种调制可以通过一个与图 5.12 类似的星座图来表示，其 8 个点均匀分布在以原点为中心的圆周上。这种 8-PSK 方案实际上已经用在一种称为 EDGE 的提高比特率的 GSM 版本上。这种方案在第 10 章第三代蜂窝系统中会讨论到。

上面的第 2 种调制形式，在设计对应于 3 比特输入序列的 8 个信号时，令信号的幅度和相位都不同。信号表达式(5.7)和式(5.7a)仍适用。8-PSK 和 QPSK 的惟一不同是信号的星座图是由包含 8 个可能的幅度组(a_i, b_i)值的映射点组成的，每个点可以定义成一个复数，在星座图平面中点的实部为 a_i，虚部为 b_i。更一般地，设 k 个连续比特被存储，则 $M=2^k$ 可能传输 k-比特序列。每个这样的序列还是可以映射成为式(5.7)那样的形式，则 M 个不同的信号 a_i 和 b_i 值必须不同。这种调制技术称为正交幅度调制或 QAM。这些信号对应于二维信号星座图中的 M 个点。QPSK 和 8-PSK 显然是这种调制的特例，其信号点都在一个圆周上。图 5.13 表示了一个 16 点星座映射的具体例子。在这种情况中，每 T s 传输式(5.7)形式的 16 种可能信号中的 1 种。在这种情况下 4 个连续的比特将被存储来产生相应的 16-QAM 信号。由于

16种可能信号的长度是PSK信号的4倍,传输这种信号序列所需的带宽就是传输PSK信号带宽的1/4。反之,给定具体的信道带宽,就能够以原来4倍的比特率传输。更一般地,存储 k 个连续比特,波特长度为原来的 k 倍,允许给定带宽上的比特率为原来的 k 倍。除了8-PSK以外,QAM没有广泛地被蜂窝无线系统使用,原因是它对合成的传输信号的确带来了幅度的变化。它被应用在无线局域网中(LANs),这将在第12章中介绍。在5.4节中,对正交频分复用即OFDM的讨论还会提到QAM,OFDM是一种使用在无线网络高速数据传输上的技术。更加深入地探讨这种技术将非常有趣。为了简单起见,我们通过其在电话调制解调器中的应用来研究这种技术。QAM是28.8 Kbps的电话线调制解调器的基础,这是在现在普遍使用的56 Kbps的调制解调器之前所使用的(56 Kbps的调制解调器也使用了OFDM,但是它与所讨论的OFDM应用这个中心偏离太远)。

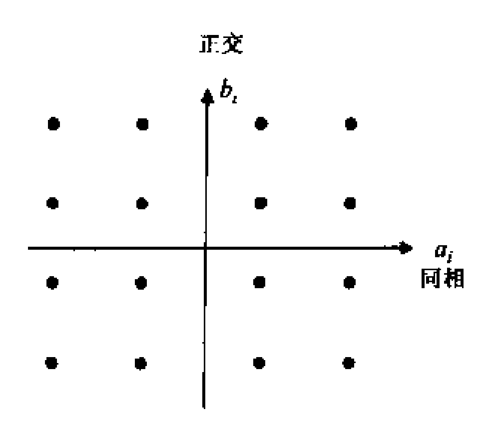


图5.13 16-QAM信号星座图

通过下面简单的例子来介绍电话调制解调器中的QAM。设家用或商用的铜电话线可以传输频率范围从300 Hz到3000 Hz的信号。QAM用来在信道中传输二进制信息,因此调制解调器是QAM类型的。带宽的中心频率选择为 $f_0 = 1650$ Hz。这个带宽与无线系统中使用的更高的RF带宽等同。在图5.14中给出了一个图表。使用12.5%的滚降成形,也就是说,$r = 0.125$,因此图5.14中的频率 f_c 是1200 Hz,而频率 f_x 为150 Hz。传输带宽也因此为2700 Hz。回顾图5.9,在二进制通信的情况下,有 $1/R = \pi/\omega_c = 1/2f_c$ Hz或 $R = 2f_c$。在QAM传输的情况下,符号率为 $1/T = R/k$,以波特为单位,与二进制情况下的 R 意义相同,因此有 $1/T = 2f_c$。具体地,在图5.14的电话例子中,$1/T = 2400$ 波特。考虑两种情况:第1种情况,$k = 4$ 比特被存储,产生 $2^4 = 16$ 可能的QAM信号,如式(5.7)中表示。实际的传输器或调制器是图5.11(a)的形式,不同的是4个连续的比特被存储而不是图中所示的2个。星座图如图5.13所示。这种方案有效的比特率是 $R = 4/T = 9600$ bps。第2种情况,令12个连续比特被存储,则在每个间隔 $T = 1/2400$,$2^{12} = 4096$ 种可能的QAM信号中的一种被传输,传输哪一种取决于进入调制器的特定的12比特的序列。星座图有4096个点,对应于相同带宽2700 Hz电话线的2400波特符号率的可能比特率,现在为28.8 Kbps。这是在出现56 Kbps之前计算机所使用的调制解调器的比特率。

在这个简单的介绍之后,回到本节先前的话题,数字蜂窝移动领域(下一节在介绍OFDM时也将简单地回到QAM问题)。主要看一下QPSK,这种调制没有幅度变化,从图5.12中可以明显地看出来。前面提到,这种调制方案使在带宽受限的无线信道上的比特传输率加倍,这也

是传输环境所要求的特性。系统仍然存在在 T s 间隔内当从一个 QPSK 信号转换到另一个时产生的相位陡峭变化到 180°(π)的问题。这个问题在 PSK 中当 a_i 和 b_i 改变符号时也同样存在。察看图 5.12 的两个星座图,可以看到这是显然的。注意到如果系数中的一个变化,相位则变化 90°(π/2)。一种能够在带宽受限的无线信道上减少最大相位变化并能够保持幅度大致不变的传输技术称为偏移 QPSK(OQPSK)或者最小频移键控(MSK)。这种方案是现在一些蜂窝无线系统所选择的调制方案的基础(Pasupathy,1979)①。在 OQPSK 中通过一种或两种(正交)载波的调制在时间上有 1 个比特的偏移($T/2 = 1/R$ s)。图 5.15 中有一个模块示意图。图中两条输入线左边的系统与图 5.11(a)中的系统相同。

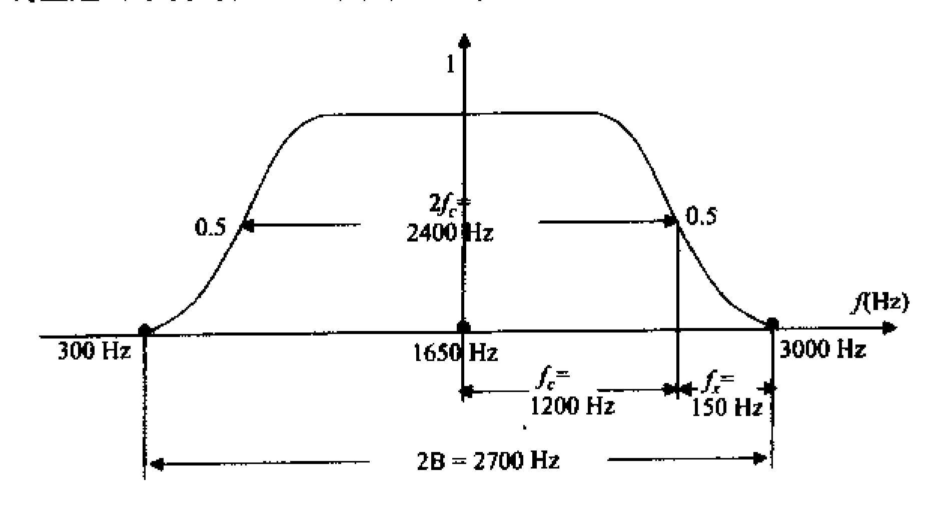


图 5.14 电话调制解调带宽

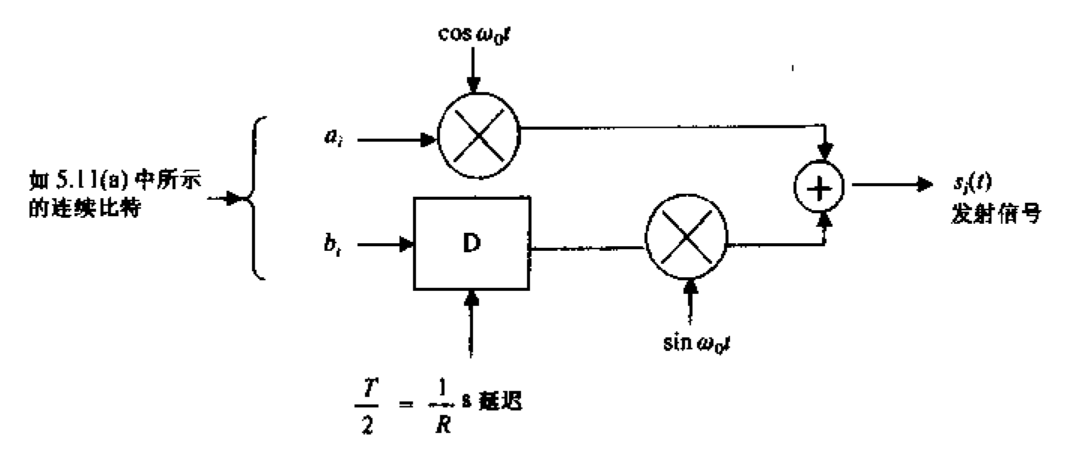


图 5.15 OQPSK 模块示意图;T = 波特或符号间隔

MSK 与之类似,只是其成形或乘以正弦函数 $\cos\pi t/T$ 的操作在图 5.15 中的乘法(调制)以及延迟函数之前被执行,如图 5.16 所示。注意到图中出现的缩写词"*lpf*"代表通过 $\cos\pi t/T$ 成形,即代表低通滤波(GMSK 使用高斯滤波,在后面将讨论到)。两条输入线左边的系统与图 5.11(a)中的系统相同。(a_i,b_i)还是代表两个连续信息比特到 4 种可能组合的映射,其例子在表 5.1 中给出。

MSK 的波形可以写成如下的形式

① Pasupathy, S. 1979. "Minimum-shift keying: a spectrally efficient modulation," *Communications Magazine*, 17, 4(July), 14-22.

$$s_i(t)=a_i\cos\frac{\pi t}{T}\cos\omega_0 t+b_i\sin\frac{\pi t}{T}\sin\omega_0 t\quad 0\leqslant t\leqslant T \tag{5.8}$$

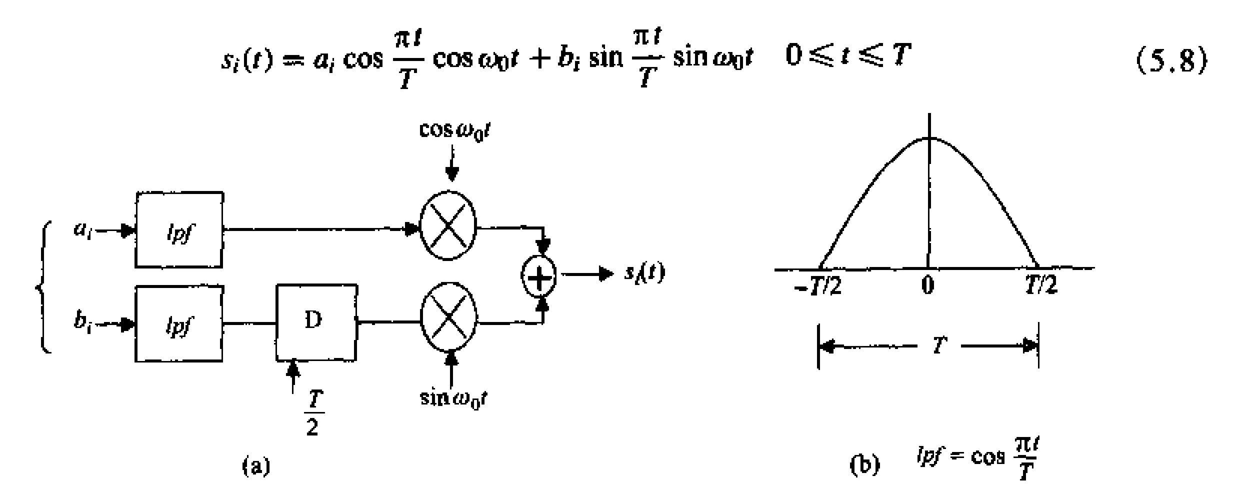


图 5.16 MSK 调制

可见,MSK 的信号与式(5.7)的 QPSK 信号类似,只是成形函数 $\sin\pi t/T$ 乘以正交载波项与成形函数 $\cos\pi t/T$ 乘以同相载波项的相位相同。$\sin\pi t/T$ 的出现是由于调制器正交项低通滤波成形的延迟 $T/2=1/R$。注意到系数 a_i 和 b_i 每隔 $T=2/R$ s 发生变化。通过三角运算,式(5.8)可以写成下面等同的形式:

$$s_i(t)=\cos\left(\omega_0 t-\frac{a_i b_i\pi t}{T}+\theta\right)$$
$$\theta=0,a_i=1;\theta=\pi,a_i=-1 \tag{5.9}$$

从这种形式的 MSK 信号 $s_i(t)$ 可以看到信号具有要求的恒定包络。原因是 $T/2$ 延迟分离了 a_i 和 b_i 的影响(如图 5.16 所示),每次只有一个系数可以变化。因此,最大的可能相位变化是 $\pi/2$,而不是 QPSK 中的最大相位变化 π。

为什么有最小频移键控这个名称呢? 这可以通过把 MSK 的表达式写成另一种形式来解释

$$s_i(t)=\cos(\omega_0 t\pm\Delta\omega t+\theta)$$
$$\Delta\omega=\pi/T,\quad \Delta f=1/2T=R/4,\quad a_i b_i=\mp 1 \tag{5.10}$$

这个公式是频移键控(FSK)的形式(见 5.1 节),频率偏差为 Δf,单位是 Hz,频率间隔用 Hz 表示为 $2\Delta f=1/T=R/2$,R 是信息比特率。在这个意义下出现了最小频移键控这个名称。为了证明这一点,来看无线信道另一端的接收端对信息比特序列的恢复。到现在为止还没有讨论过称为检测的过程,只是在 5.1 节对 PSK 的讨论中提到过,在这种情况下需要一个相位的参考。为了简化讨论,在式(5.10)中令 $\omega_1=\omega_0-\Delta\omega$。对于 $a_ib_i=+1$ 的情况,令 $\omega_2=\omega_0+\Delta\omega$,对于 $a_ib_i=-1$ 的情况,暂时令 $\Delta\omega$ 为任意值。图 5.17 表示了一个对 MSK 检测也适用的同步 FSK 检测器。同步检测器这个名称意味着乘以图 5.17 中两个分支的接收 $s_i(t)$ 信号,其本地产生的正弦信号的相位和频率与相应的传输信号是相同的或锁定的。标记"lpf"的模块是低通滤波器,其频率响应比高频 MSK 信号的频率低得多(或者一般而言的 FSK 信号)。可以把这个设备近似成为一个积分器,可提供最简单的低通的平滑操作。判决点是那些在信号每隔 $T/2$ s 被采样的点。在这些采样点根据 a_i 或 b_i 的值进行判决,随着传输器产生的一个一个比特轮流更替。因此,决定了 a_i 的值,利用这个值来判定 $T/2$ s 后的 b_i 传输值;给出了 b_i,再用它来判定 a_{i+1} 的值,重复这个过程。

现在来看这些判决是如何做出的。举例,设已经得到 a_i 是 +1。在判决 b_i 时,要通过采样 $T/2$ s 后的两个判决点,考虑两种可能性。如果 $b_i = +1$, $a_ib_i = 1$,从式(5.10)中得到,传输的是 $\cos(\omega_0 - \Delta\omega)t = \cos\omega_1 t$。图 5.17 中上面的位置 1 的低通信号采样的值可以通过在 $T/2$ s 上的积分(低通滤波器输出)$2/T[\int_0^{T/2}\cos^2\omega_1 t dt]$来近似。读者可以看到,对于 $\omega_1 T \gg 1$ 这个合理的假设(为什么?),该值非常接近 1/2。考虑下面相同采样时间的判决点 2。现在来说明,在采样条件下,其值是 0。具体地,点 2 的采样值可以近似为

$$\frac{2}{T}\int_0^{T/2}\cos\omega_1 t\cos\omega_2 t\mathrm{d}t = \frac{1}{T}\int_0^{T/2}[\cos(\omega_1+\omega_2)t + \cos 2\Delta\omega t]\mathrm{d}t$$

$$\approx \sin\Delta\omega T/2\Delta\omega T$$

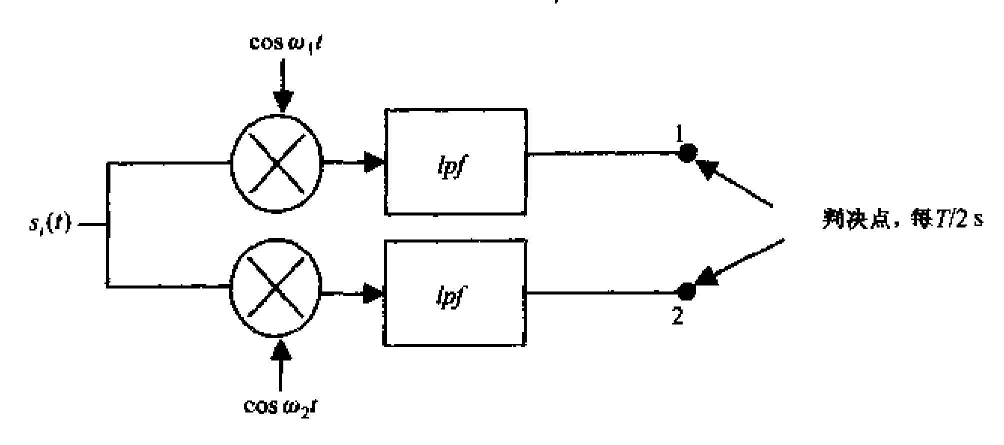


图 5.17 FSK 信号检测;惯用于检测 MSK 信号

这里用简单的三角函数来代替余弦函数的和以及偏差角度的乘积,利用了 $\omega_2 - \omega_1 = 2\Delta\omega$,并假设$(\omega_1 + \omega_2)T \gg 1$,现在我们选择 $\Delta\omega T$ 使 $\cos\Delta\omega T = 0$。这样,根据选择 $\Delta\omega T$,通过分析表明如果已经确定 a_i 为 +1,b_i 为 +1,则会在接收端的判决点 1 处产生正的信号,而在点 2 处却没有输出。读者可以自己得到这样的结果。使用相同的分析,$b_i = -1$ 会在点 2 产生信号,而在点 1 处无信号产生。类似地,如果 $a_i = -1$, $b_i = +1$ 在点 2 产生输出,而在点 1 不产生信号,则 $b_i = -1$ 时与此相反。这样给定了已知的 a_i,在接收端就可以对 b_i 值做出惟一的判断。继续这个论断,读者也可以在 b_i 已知时,对 a_i 做出惟一的确定。根据 a_i 和 b_i 序列,传输的信息相关比特流就可以在接收端轮流重建恢复。但是,注意到正确地做出判断需要有一个判决点具有非零的信号采样。这种情况发生的条件是前面已经提到的 $\Delta\omega T = 0$。满足这个条件的最小频率偏差 $\Delta\omega$ 是 $\Delta\omega T = \pi$ 或 $\Delta f = \Delta\omega/2\pi = 1/2T = R/4$,这正是式(5.10)中的 MSK 条件。总结一下,MSK 调制可看成是一种 FSK 调制方案,它使用了允许在接收端正确检测传输信号流的最小频率偏移(这并不是说传输的信号总是正确的。系统中的噪声、衰落效应、导致符号间干扰的时间抖动以及其他发送器的干扰,也会偶然引起信号检测错误的发生,这是一个设计优良的系统所不能承受的)。

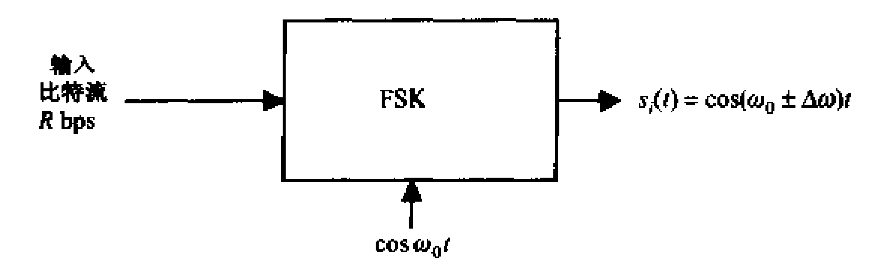
现在来看 GSM 移动无线系统中使用的高斯 MSK,即 GMSK。这种调制方案推广了 MSK,使其包括了高斯成形。图 5.18 表示了这个过程。在图 5.18(a)中表示了等同于 MSK 的 FSK,在上面的式(5.10)中已经给出。在 GMSK 中,FSK 调制器之前有一个低通高斯成形滤波器 $h(t)$,如图 5.18(b)所示。这个成形滤波器具有下面式(5.11)的形式,这也是它名字的由来。

$$h(t) = \mathrm{e}^{-t^2/2\sigma^2} \tag{5.11}$$

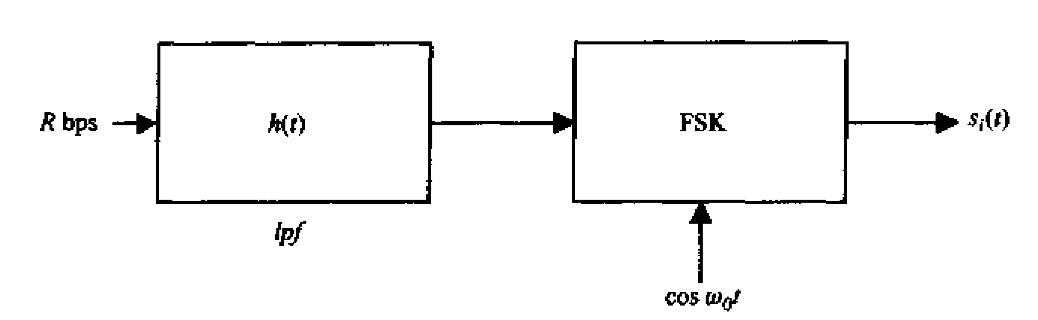
成形将运用到每个输入信息比特,如图 5.18(b)所示。参数 σ 控制进入调制器的合成成形二进制脉冲的宽度。增加 σ 就会增大脉冲延展进入相邻的比特间隔,增加符号间的干扰。与正弦滚降成形中选择特殊的成形函数使符号间干扰为零的情况不同,除了时域抖动效应外,这里采用受控的符号间干扰。原因是通过增大成形后的二进制脉冲宽度,传输带宽会减小,这表示了众所周知的脉冲宽度和带宽之间的反向关系(Schwartz, 1990)。这个相反的关系通过计算 $h(t)$的傅里叶变换即其频谱来证明。容易看出,$H(\omega)$也是高斯成形的形式,由下面的式(5.12)给出

$$H(\omega) = C\mathrm{e}^{-\sigma^2\omega^2/2} \tag{5.12}$$

C 是常数。3 dB 带宽,也就是当 $H(\omega)$是它峰值的 $1/\sqrt{2} = 0.707$ 时用 Hz 表示的带宽,容易得出 $B = 0.133/\sigma$。可以看出,脉冲宽度 σ 和带宽 B 之间的反向关系。成形函数 $h(t)$和 $H(\omega)$如图 5.19 所示。

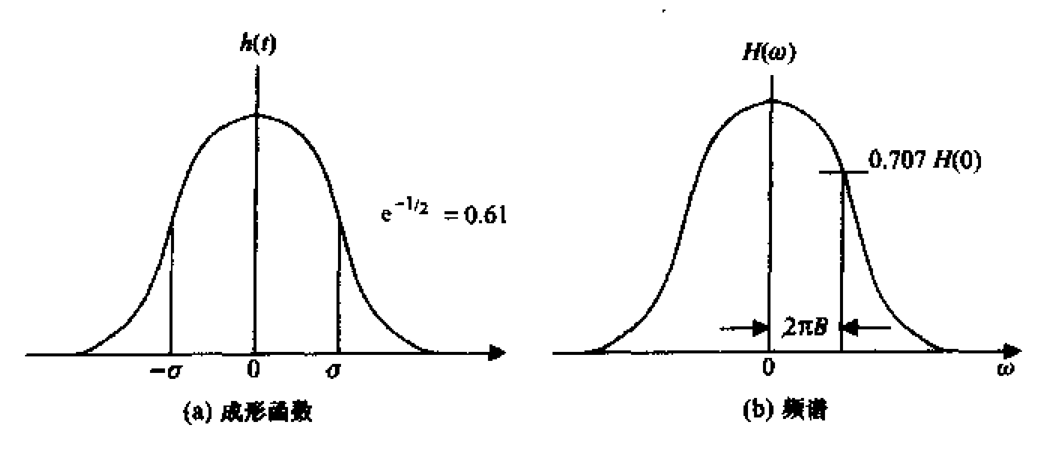


(a) 与 MSK 等同的 FSK



(b) $h(t)$: 高斯成形 MSK

图 5.18 高斯 MSK 调制



(a) 成形函数 (b) 频谱

图 5.19 高斯成形

成形函数 $h(t)$ 的另一种表达形式可以用带宽 B 所表示的等同式代替 σ,写成下面的形式

$$h(t) = e^{-t^2/2\sigma^2} = e^{-(2\pi B)^2 t^2/2\ln 2} \tag{5.11a}$$

可见当 $\sigma \to 0$ 或 $B \to \infty$ 时,成形效果就会消失,在调制器的输出端可以得到原始的 MSK 信号。

图 5.18(b)中成形函数对 FSK 调制器输入的作用可以通过用一个未成形的方形代替二进制脉冲来得出。设脉冲代表通过(低通)成形器而没有经过调制器的二进制数为 1。脉冲的宽度是 $1/R = T/2 \equiv D$,激发了图 5.18(b)的高斯滤波器。所需的输出 $g(t)$,也是 FSK(MSK)调制器的输入,可以通过两个时间函数的卷积来得到。合成的时间函数由式(5.13)给出

$$g(t) = \text{erfc}[2\pi B(t-D)/\sqrt{\ln 2}] - \text{erfc}[2\pi B(t+D)/\sqrt{\ln 2}] \tag{5.13}$$

式(5.13)中的互补错误函数 $\text{erfc}(x)$ 是高斯函数的积分

$$\text{erfc}(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-y^2} dy \tag{5.14}$$

高斯成形的作用可以通过将成形的函数 $g(t)$ 作为标准化带宽 $B_N \equiv BD = BT/2 = B/R$ 的函数来得出(Steele,1992:535)[①]。当 B_N 减小而输入比特率 R 固定时,则带宽减小,$g(t)$ 就更加延伸到相邻的二进制间隔中,增加了符号间干扰。而合成的频谱,即 $g(t)$ 的傅里叶变换,或者等同于式(5.12)的 $H(\omega)$ 乘以方形二进制输入脉冲的 $\sin x/x$ 频谱,与 MSK 的频谱相比就减少了频谱过载(Steele,1992:540,541)。因此 GMSK 比普通的 MSK 所需的传输带宽减小,代价是符号间的干扰增大了。一个 B_N 的较好的折中值是 $B_N = 0.3$。这个值提供了传输带宽的一个充分的减小,且并不带来太大的符号间干扰。这个值实际上就是 GSM 系统在使用 GMSK 调制中采用的值。

第二代北美 TDMA 蜂窝系统 IS-136(先前的 IS-54)或 D-AMPS 的设计者选择的调制方案是 $\pi/4$-DQPSK(差分 QPSK)。这个方案的信号星座图使用了 8 个不同的相位及恒定包络。4 个不同的 (a_i, b_i) 值与星座图相位相应的差分变化相对应(Gibson,1996:Ch.33)[②]。给定系统星座图中的一点,设在 T s 间隔 n 的结尾,系统可以移动 4 个不同的点,这取决于下一个 T s 间隔 $n+1$ 出现 4 个 (a_i, b_i) 值中的哪一个。表 5.2 给出了 (a_i, b_i) 对差分相位变化 $\Delta\theta(n) = \theta(n+1) - \theta(n)$ 的映射。8 点星座图的图示以及差分相位变化的例子在图 5.20 中给出。在这个例子中,系统在间隔 n 末尾的相位点设为 $3\pi/4$,图中已经标出。下一个 T s 间隔的 $(a_i, b_i) = (-1, -1)$ 使相位变化 $-3\pi/4$,旋转相位点到 $\theta(n+1) = 0$。对应于表 5.2 中所表示的其他 3 个可能的旋转在图 5.20 中指明。这样,传输信号与前面式(5.7a)中的信号类似,是包络恒定相位变化类型,使用 8 个可能的相位位置。与 QPSK 的情况一样,在连续的传输信号中信号点之间的转换会导致幅度的变化。但是,这些变化像 8-PSK 那样,相比 QPSK 有所减小。

在这种情况中如何完成接收端的检测呢?显然,相位检测是必需的,因为要恢复的信息承载在所接收信号的相位上。前面讲到在相移键控和 FSK 类型的调制中,接收端必须锁定在传输信号的相位和频率上。为了保证正确的检测,这样的技术需要接收端能够正确地跟踪传输信号完全的相位情况。而在差分相位调制的方式中,这种完整的相位信息则不是必需的。差

① Steele, R. ed. 1992. *Mobile Radio Communications*, London, Pentech Press; New York, IEEE Press.

② Gibson, J. D. ed. 1996. *The Mobile Communication Handbook*, Boca Ration, FL, CRC Press/IEEE Press.

分相移判决只需要在每个 T s 间隔来决定发送器产生的(a_i, b_i)值，通过这样的方式，相应的二进制信息序列就可以被重建和恢复。其实，一个接一个的传输信号为后面的信号提供了必要的相位参考。这还可以显著地简化相位检测的过程，但是必须对检测和纠正错误的相位判决进行事先规定，否则相位错误就会无限期的存在。

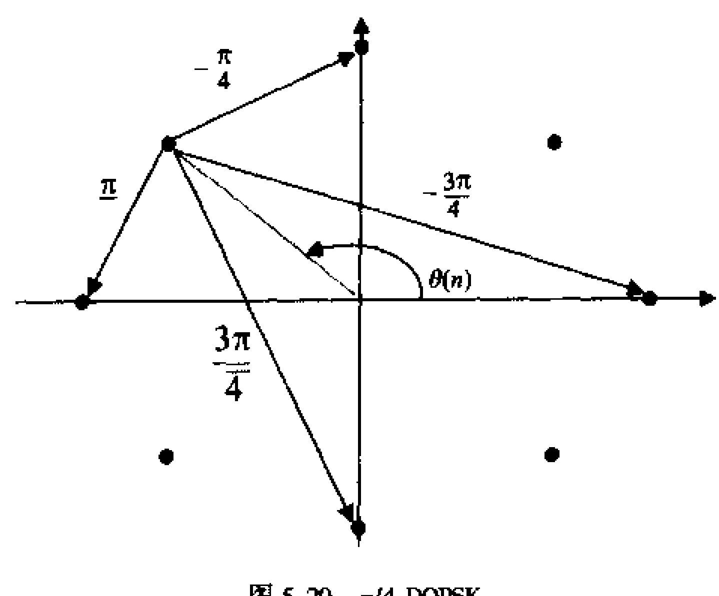


图 5.20 π/4 DQPSK

5.4 正交频分复用

这一节介绍正交频分复用(OFDM)。在本章的简介中已经提到，OFDM 现在应用在基于铜电话接入线的高速 DSL 调制解调器中。它也作为 IEEE 标准 802.11g 和 802.11a 的一部分被标准化，用在无线局域网(WLAN)中 54 Mbps 的高速数据传输上。这些应用是第 12 章对 WLAN 讨论的一部分。OFDM 也被建议用于先进蜂窝系统中。

OFDM 对无线系统的应用是基于无线系统对更高比特率、更高数据传输带宽的要求而来的。在第 2 章中看到当传输带宽增大时，就会产生频率选择性衰落和随之发生的信号失真。在数字传输中，产生符号间干扰，连续的数字符号交叠到相邻的符号间隔当中。OFDM 通过把信号传输频谱分成狭窄的片段，并在这些片段中平行地传输信号从而减弱了这个效应。如果这些频谱段足够狭窄，平坦或非频率选择性衰落就会产生，每段上的传输信号将会被无失真地接收。在这个意义上，OFDM 对无线应用具有特殊的重要性。

具体分析，设二进制数位以比特率 R bps 传输。需要的传输带宽 B 由式(5.6a)得到，为 $R(1+r)$，r 是奈奎斯特滚降因子。设有 N 个这样的比特序列，其间隔为 $T_S = N/R$，称这个间隔为 OFDM 符号间隔。这样就实施了串-并变换，用 N 个比特序列的每一个分别来调制一个载波。所有 N 个调制后的载波信号在 T_S 长的间隔上同时传输，这个过程就是 OFDM 技术的本质所在。图 5.21(a)描绘了这个简单的 OFDM 产生机制。参数 $a_k, 1 \leq k \leq N$，代表连续的比特，频率 $f_k, 1 \leq k \leq N$，代表 N 个平行传输的载波频率。图 5.12(b)表示了合成的 OFDM 频谱。为了使不同的频率载波互相正交，需要满足载波间的频率间隔 Δf 等于 $1/T_S$。可得 $B = N\Delta f$，B 是图 5.12(b)中表示的传输带宽。Δf 也就是 N 个平行频道中每个的带宽。

很明显，这个过程使每个平行传输的信号的传输带宽变成了原来的 $1/N$。如果 N 足够

大，平坦衰落而不是频率选择性衰落就会出现在每个使用的频道上，这样就能够不通过串-并变换来克服任何的频率选择性衰落（注意到这个先存储比特序列然后用 N 个载波传输的过程与 QAM 不同，在 QAM 中只使用一个载波）。需要怎样的 N 值或者 Δf 值呢？回顾第 2 章，平坦衰落的发生需要传输带宽 B 和时延扩展 τ_{av} 满足下面的关系：

$$\tau_{av} < 1/2\pi B \tag{5.15}$$

因此，必须有 $\Delta f < 1/2\pi\tau_{av}$。用时间等效，利用 $\Delta f = 1/T_S$，则必须有

$$T_S = N/R \gg \tau_{av} \tag{5.16}$$

简单地说，为了避免符号间干扰，OFDM 的符号间隔必须比时延扩展大得多，这是直观上得到的结论。

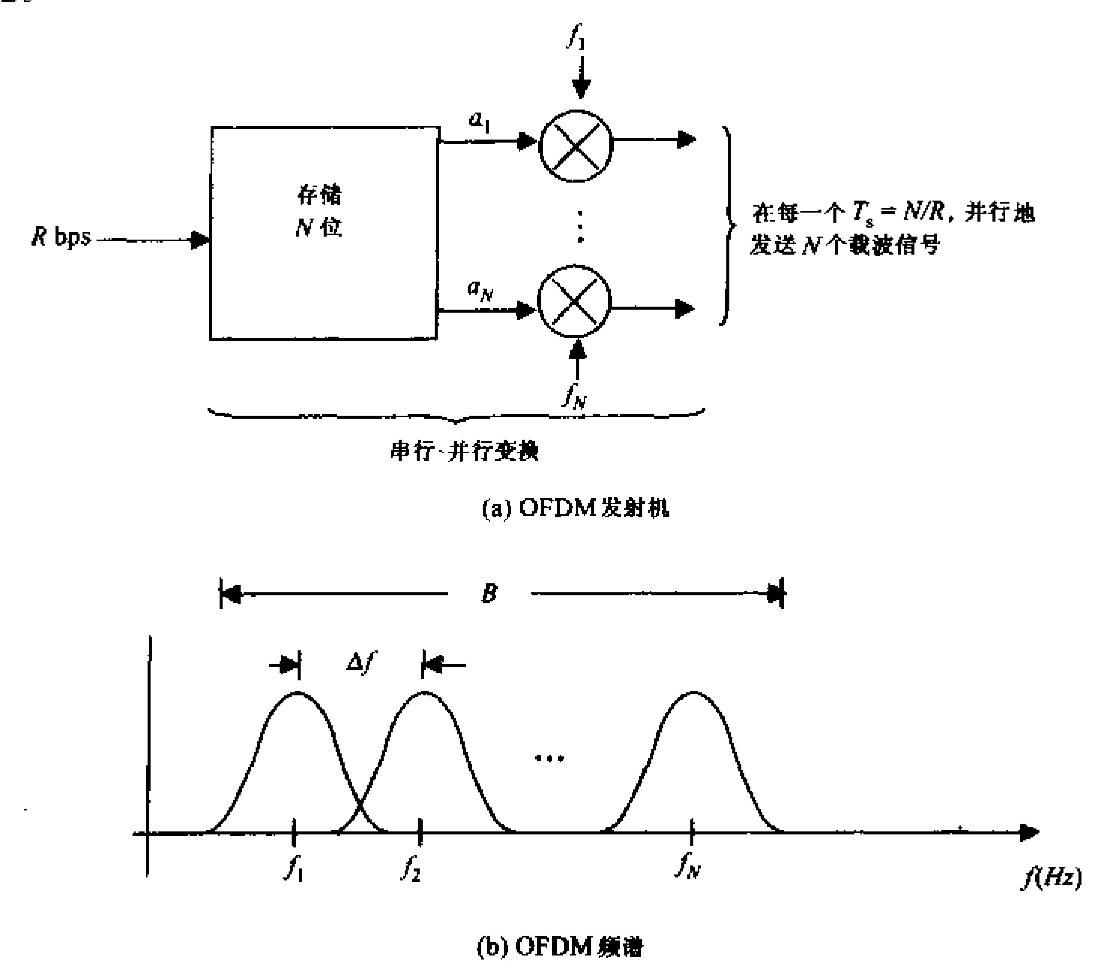


图 5.21 正交频分复用（OFDM）

看一些实例。第 1 个例子，设传输带宽是 1 MHz，在此信道上传输的比特率是 800 Kbps（这里的奈奎斯特滚降因子为 0.25）。时延扩展为 1 μs，会引起频率选择性衰落，同时产生符号间串扰。利用 10 个间隔为 100 kHz 的载波的 OFDM，符号间串扰就会减弱，在 10 个平行的 100 kHz 宽的信道上干扰就是平坦衰落（这些 OFDM 载波通常称为子载波）。

第 2 个例子，考虑一个 20 MHz 宽的信道，这是使用在高数据率无线局域网的 IEEE 802.11g 和 802.11a 标准上的。如果不使用 OFDM，大于 10 ns 的时延扩展将会导致频率选择性衰落。而从第 2 章知道，在不同通信环境中的时延扩展，包括在微型和室内环境中的时延扩

展，都远大于这个值。因此，为了获得更大带宽的系统，我们就是需要像 OFDM 这样的技术。载波间隔，即子载波的带宽为 312.5 kHz，在第 12 章中可以看到这个值实际上已经在 20 MHz 宽的无线局域网系统中使用。通过 312.5 kHz 子载波带宽的 OFDM，大于 0.5 μs 的时延扩展会带来符号间串扰，这比没有使用 OFDM 的情况有了实质上的改善。

使用多载波传输的 OFDM 在实际执行的时候会产生一定的问题。现在来证明通过分析这个问题可以通过离散傅里叶变换来解决。考虑图 5.21 的 N 个平行的输出信号在 T_S 长符号间隔上的传输。记这些信号的和，即传输的所有信号为 $v(t)$。定义第 k 个载波频率 f_k 为 $f_c+k\Delta f, 0\leqslant k\leqslant N-1$。重新定义子载波频率的命名，令 N 个平行子载波最低的频率为 f_c，其他的都在这个值上以 Δf 为间隔分布。第 k 个载波可写成 $2\pi(f_c+k\Delta f)t$，$v(t)$ 可以写成

$$
\begin{aligned}
v(t) &= \mathrm{Re}\left[\sum_{k=0}^{N-1} a_k \mathrm{e}^{\mathrm{j}2\pi(f_c+k\Delta f)t}\right] \\
&= \mathrm{Re}\left[\mathrm{e}^{\mathrm{j}2\pi f_c t}\sum_{k=0}^{N-1} a_k \mathrm{e}^{\mathrm{j}2\pi k\Delta f t}\right] \\
&= \mathrm{Re}\left[\mathrm{e}^{\mathrm{j}2\pi f_c t}a(t)\right]
\end{aligned}
\tag{5.17}
$$

其中，$a(t)\equiv\sum_{k=0}^{N-1}a_k\mathrm{e}^{\mathrm{j}2\pi k\Delta f t}$；Re[ ]代表实部。

以 T_S/N 为间隔采样 $a(t)$，也就是说，速率是每秒 R 个采样。若用采样函数 $a(n)$ 代替函数 $a(t)$，用 nT_S/N 代替 t，则 $n=0\cdots N-1$。由 $\Delta fT_S=1$，可以把 $a(n)$ 写成

$$
a(n)=\sum_{k=0}^{N-1}a_k\mathrm{e}^{\mathrm{j}2\pi kn/N}\qquad n=0\cdots N-1 \tag{5.18}
$$

可见式(5.18)正好是离散傅里叶反变换的形式(Oppenheim 和 Willsky，1997)[①]，并可以容易的通过快速傅里叶变换(FFT)来计算。只利用单个载波频率 f_c 的传输代替图 5.21 所示的和式(5.17)中数学表示的 N 个平行载波的传输，OFDM 过程可以由 FFT 计算来代替。这个等同的过程如图 5.22 所示。在第 12 章中可以看到，这就是先前用在高速无线局域网标准中的方法。在接收端，每符号间隔 T_S 对接收到的调制载波信号进行相反的过程：计算离散傅里叶变换，可以恢复 N 个系数 a_k，$k=0\cdots N-1$ 以及进行并-串行变化来产生需要的输出比特流。在进行 FFT 计算时，N 选择为 2 的倍数对计算更有利。

这个通过使用离散傅里叶反变换的二进制输入采样来产生 OFDM 信号的过程可以被推广到 QAM 上的信号采样，允许在一定带宽上传输更高比特率的信号。具体地说，设 QAM 首先用来减小传输给定输入比特流所需的带宽。令 $K<N$ 个连续的二进制数字被存储，产生 2^K 个可能的 QAM 信号。在对 QAM 的讨论中可以知道，这种调制的星座图如图 5.13 所示。我们看到星座图中代表传输的 2^K 个可能信号的每个点，都可以表示成一个复数。具体地，设第 k 个这样的数字记为 $\boldsymbol{a}_k$。在 T_s 的符号间隔上，用连续的 K 比特组进行 QAM 调制，存储合成的 N 个连续的复数 $\boldsymbol{a}_k$，$k=0\cdots N-1$，可以通过串并变换用不同的子载波对 N 个复数进行传输。这样就得到了 OFDM 的更一般的形式。这个在 OFDM 之前首先用 QAM 调制信号的过程对应于将

① Oppenheim, A. V. and A. S. Willsky. 1997. *Signals and Systems*, 2nd edn, Upper Saddle River, NJ, Prentice-Hall.

式(5.17)一般化，即用复数 a_k 来代替那里的二进制系数 a_k。同样，通过采样合成的公式使离散傅里叶反变换技术成为一种执行 OFDM 操作的方法，但是，它是以 QAM 的单位间隔 K/R 来进行采样的。我们可以得到式(5.18)的离散傅里叶反变换，但用复数 a_k 代替 a_k 可包括更一般的情况。这个经修改的更加一般的形式如式(5.19)所示

$$a(n)=\sum_{k=0}^{N-1}a_k e^{j2\pi kn/N} \tag{5.19}$$

可以利用快速傅里叶变换技术来计算式(5.19)的离散傅里叶反变换。传输信号 $v(t)$ 是先前式(5.17)的形式，但是用式(5.19)更一般的形式代替了式(5.18)。这个在执行串-并变换之前融合 QAM 的更加一般的 OFDM 系统如图 5.23 所示。IFFT 模块输出的 I 和 Q 信号代表复杂 IFFT 操作的同相和正交部分。

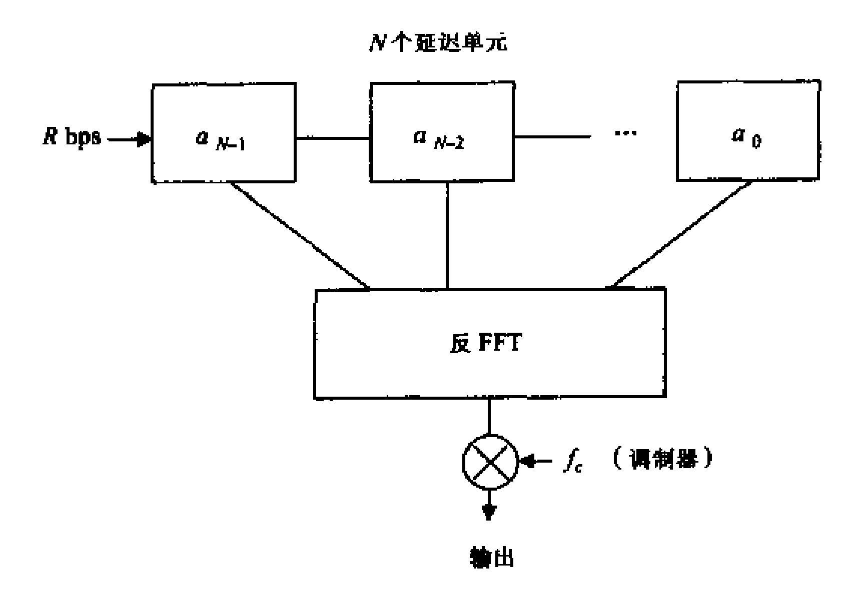


图 5.22　OFDM 信号的等效生成

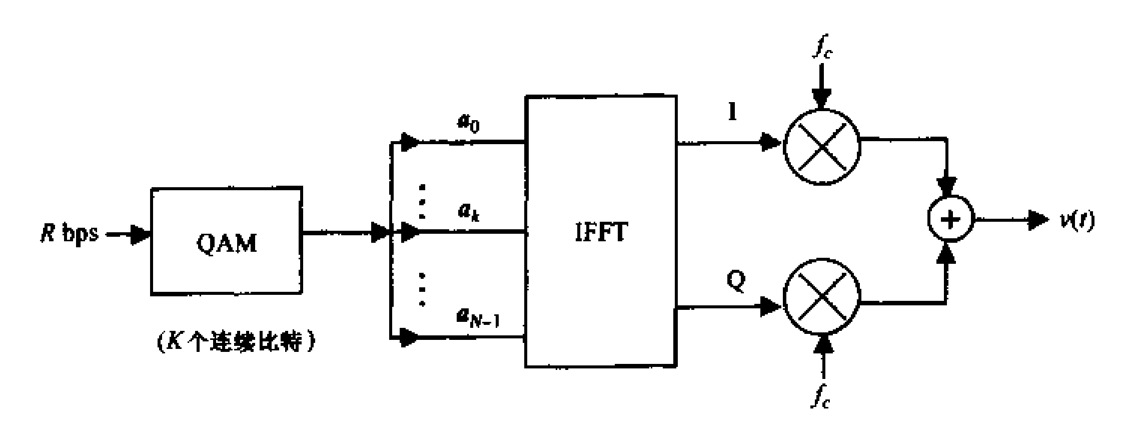


图 5.23　加入 QAM 时 OFDM 的输出

通过一些融合 QAM 的 OFDM 系统的例子来总结这一节。

例 1

设传输带宽是 1 MHz，使用的滚降因子是 $r=0.25$，则数据传输率为 800 Kbps。对时延扩展为 0.2 μs 或更多来说，频率选择性衰落以及随之而来的符号间干扰就会发生。使用 $N=16$ 个等同间隔子载波的

OFDM。子载波间隔 Δf 为 67 kHz,而 OFDM 符号间隔为 $T_S = 1/\Delta f = 15$。小于等于 3 μs 的时延扩展则不会产生符号间干扰。如果 OFDM 系统使用 16-QAM 调制,即通过存储 4 个连续比特来产生 QAM 星座图的 16 种可能的复数点,则对于相同的时延扩展数目能够以 3.2 Mbps 的速率传输。

例 2

对于同一个 1 MHz 的例子,设需要传输 4.8 Mbps 的数据率,并对 25 μs 的时延扩展不存在符号间串扰,则必须首先使用 64-QAM 来减小符号率到信道所允许的 800 k 波特。为了适应 25 μs 的时延扩展,选择至少其 5 倍的最大时延扩展,即 125 μs 作为 OFDM 的符号间隔 T_S。选择 $N = 128$ 个子载波,则子载波之间的间隔为 $\Delta f = 7.8$ kHz,符号间隔为 $T_S = 1/\Delta f = 128$ μs。

在第 12 章对高数据率 WLAN 标准 IEEE 802.11g 和 802.11 的讨论中,我们将会进一步对 OFDM 技术进行探讨。

习题

5.1 (a) 为什么在移动通信系统的无线信道上更倾向于使用恒包络调制?
(b) 在(a)的基础上解释为什么第一代 AMPS 使用的是 FM 而不是 AM?
(c) 参考(a),为什么数字移动无线系统避免使用 QAM 的一些多进制幅度调制形式,尽管这些形式可以在信道上获得更高的比特率?
(d) 为什么在移动无线系统的设计中带宽和功率是重要的考虑因素?

5.2 (a) QPSK 用来在载波频率是 871.5 MHz 的无线连接上传输数据。这个频率的可用传输带宽是 1.25 MHz。使用正弦滚降成形。计算对应于下列滚降因子的数据率:$r = 0.25, 0.5, 1$。如果使用的是 PSK,则相应的速率是多少?对于不同的载波频率,结果会变化吗?请解释。
(b) 对于中心是相同载波频率的 30 kHz 带宽的情况,重复(a)的问题。同样,对于不同的载波频率,结果会变化吗?

5.3 (a) 选择图 5.11 中应用于 QPSK 调制器的 10 个或更多比特数的输入序列。用表 5.1 将它们映射到 5 个合适的 QPSK 信号组上。用式(5.7)的信号表达来描述输出的 QPSK 信号。为简单起见,忽略成形函数。设载波频率是 $1/T$ 的倍数。标注出现 π 或 $\pi/2$ 相位变化的那些时刻,使其与输入比特相互关联。
(b) 对图 5.15 的 OQPSK 调制器重复(a)的问题。这种情况的最大相位变化是多少?

5.4 (a) 针对 QPSK,解释对于给定的二进制序列,当式(5.7)中的系数 a_i 和 b_i 有一个发生变化,以及两个都发生变化时,对相位变化所造成的影响。
(b) 证明对于(a)中相同的序列,OQPSK 和 QPSK 的区别是什么?其中,是什么使得 OQPSK 比 QPSK 更加适合移动通信系统?
(c) 对(a)中的同一个序列比较 MSK 和 QPSK。

5.5 (a) MSK 调制通过式(5.8)及图 5.16 来定义,证明式(5.8)可以通过式(5.9)和式(5.10)的等同形式来表示。
(b) 通过例子解释文中的论断,即对于 MSK 系数 a_i 和 b_i 每次只有一个发生变化。为什么最大的相位变化只有 $\pi/2$ 呢?

5.6 考虑图5.16中表示的MSK调制器。画出任意选择的至少10个输入比特序列的输出MSK信号图。证明信号确实具有恒包络。在比特变化点存在相位不连续点吗？试解释。

5.7 (a) 式(5.10)对MSK检测的分析表明，如果 $f = R/4$，则连续的 a_i 和 b_i 的值可以通过图5.17的FSK检测器惟一恢复。给出这个分析过程，补充没有指明的分析细节，并证明论断的正确性。

(b) 继续(a)中对图5.17的FSK检测器的讨论，证明文中指出的对所有 $a_i(b_i)$ 的可能值，检测器都能够在 $T/2$ 间隔之后惟一正确地产生被传输的 $b_i(a_i)$ 值。

5.8 式(5.13)表示了用高斯滤波器对宽度为 D 的二进制脉冲响应进行成形的作用，该高斯滤波器的脉冲响应在式(5.11)给出，其频谱用式(5.12)表示。对不同的 $B_N = BD$ 值，画出式(5.13)作为时间的函数图，包括特定的值0.3以及此值上下的值。正如文中所指出的，减小带宽的结果是引入了更多的符号间串扰。需要使用式(5.14)中所定义的互补错误函数 erfc(x)软件包或查表。

5.9 考虑正文中描述的π/4 DQPSK。系统的初始相位设0。设进入的比特流为0 0 1 0 1 1 1 0 0 1 1 1，顺序为从左到右。在图5.20的星座图中随着输入比特一个接一个被处理，指出调制器移动过的点。

5.10 可用的传输带宽为2 MHz。使用奈奎斯特滚降成形传输数据。

(a) 得到在该信道上的比特率，使用PSK调制，滚降因子分别为(1)0.2，(2)0.25和(3)0.5。

(b) 使用的奈奎斯特成形因子为0.25，该信道上的传输比特率为6.4 Mbps，表明这是如何得到的。对9.6 Mbps重复该问题。

5.11 习题5.10的信道用在比特率是6.4 Mbps的无线数据传输上。

(a) 估计可能容忍的最大时延扩展。

(b) 需要满足的时延扩展为8 μs。使用OFDM。找到合适的载波数目、每一个的带宽以及合成的OFDM符号间隔。画出合成的频谱图。

(c) 如果实施时使用离散傅里叶变换，画出(b)中OFDM发送器的模块图。

第6章　多址接入技术：FDMA，TDMA，CDMA系统容量比较

在第3章和第4章的FCA和DCA的讨论中，比较抽象地使用了“信道”这个词。在这一章中，我们将会使这一概念更加具体，并提供现行蜂窝系统中不同类型的“信道”的例子。信道这个词指的是一种分配给指定用户的系统资源，以便该用户通过网络与其他用户在可容忍的干扰条件下进行通信。这意味着信道之间应互相正交。蜂窝系统中最普遍使用的信道类型是频分信道、时隙信道和编码信道。这3种为蜂窝网络系统用户提供接入的方法，分别称为频分多址接入(FDMA)、时分多址接入(TDMA)和码分多址接入(CDMA)。在这一章中会描述这些不同的多址接入技术，并以3种最常用的第二代数字蜂窝网络系统作为例子。这3种系统中都使用了FDMA技术。其中，两种系统GSM和D-AMPS或IS-136，是基于TDMA的；第1种系统IS-95使用的是CDMA技术。由于FDMA是这本书中将要讨论的所有蜂窝系统的基础，包括第10章将要讨论的第三代系统，因此将先简要介绍一下FDMA的概念。后面再用单独的小节讨论TDMA和CDMA系统，并在本章结束时比较这3种系统的“信道容量”，或者说比较特定的频带中各多址接入方案所能提供的每小区的用户数。

首先来看FDMA的概念。这一概念很容易解释。它最简单的形式是把给定的频带分成不同的信道，每个子信道分配给不同的系统用户或者移动终端。这就是第一代模拟移动系统使用的方法。当然，历史上频带的子信道化方法最早用于广播式收音机和广播式电视的节目传送中。在播送过程中，每个播送站分配到一个给定的信道或者频率范围，使用例如幅度调制(AM)或频率调制(FM)的调制方法，子信道之间留有足够的保护带宽来保证它们之间只有有限的干扰，而地理上距离足够远的广播和电视站可以重复使用相同的信道而彼此不会受到干扰。第一代的模拟蜂窝系统就是在这个原理上实施的。例如，第1章中提到的北美AMPS系统，其下行链路或前向信道[即从基站(BS)到移动台(MS)方向]分配到869 MHz到894 MHz的频率范围。上行链路或反向信道的传输在824 MHz到849 MHz的频率范围上实行。这两个25 MHz的频带各被分成832个30 kHz的信道，每个移动用户分配一对信道，上行和下行各分配一个信道。这样的频率信道分配有时称为正交频率分配，它允许多用户同时使用系统。AMPS系统是单纯的FDMA系统。通过频分信道的成对组合使系统能同时在上行或下行进行双向通信的方法，称为频分双工或FDD。因此，不同方向的通信有单独频率信道的复合方案，被标记成FDMA/FDD。注意与广播模式的区别，后者只有一个通信方向，且所有接收者使用同一个信道。

下面的部分用来讨论前面所说的TDMA和CDMA，使用的例子是GSM，D-AMPS和IS-95。这些第二代系统在第8章中将有更深入具体的讨论(第三代数字系统包括使用时分多址和码分多址接入技术的系统，在第10章中将有详细的介绍)。在6.1节中讨论了TDMA系统，用GSM和D-AMPS的例子进行说明。6.2节中介绍了CDMA。6.3节到6.5节介绍了CDMA系统的系统容量，用给定小区能够容纳的用户数量这个参数来衡量。利用6.5节中介绍的性能结果，在6.6节中进行了GSM，D-AMPS和IS-95 CDMA系统的系统容量的比较。

6.1　时分多址接入技术

TDMA 系统，例如，全欧洲 GSM 系统和北美 IS-136 或 D-AMPS 系统，通过将多个用户分配到同一频率信道中，获得了比第一代 FDMA 系统更大的容量。这些第二代 TDMA 系统实际上是开关系统，它的一个例子就是现代数字电话系统。在第二代 TDMA 中，数字信号在相应信道载波频率上的重复帧结构中以一定时隙的形式进行传送。每个用户分配到每帧中一个或多个时隙，并在需要的时间内保存时隙，直到一个“通话”完成。数字信号本身采用某种调制方案传输（PSK 或 QPSK 的一种），如在第 5 章中已经介绍的那样。图 6.1 表示了这样一个每帧中包含 N 个时隙的例子。当然，为了区分帧的开始和结尾，同时也为了在帧中给时隙定界，我们还有必要进行另一些设置。如前面所提到的，每个用户分配到每帧中的一个或多个分片。这些第二代系统主要设计用来传送通话语音，就如数字有线电话系统一样，当然数据传送也开始扮演越来越重要的角色。在第 8 章中会详细介绍这些系统的声音通话传输。而第 10 章中介绍的第三代系统则主要设计用来传输数据并使用数据包转换技术，即因特网使用的那种技术。

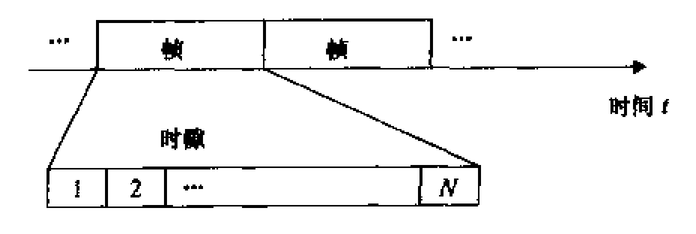


图 6.1　TDMA 时隙结构

第二代 TDMA 系统实际是 FDMA/TDMA 系统，因为系统的整个频谱首先分成了若干有特定载波频率的子频率信道或频带。然后，每个频带按某种帧结构被分为若干时隙以支持多个用户。图 6.2 表示了整个 FDMA/TDMA 的框图。在一个特定的频带中，一个“信道”对应于一个或多个时隙。在图中有 M 个频带，每个具有不同的载波频率，每帧包含 N 个时隙，分派到一个指定的载波上。图中表示了分配到一个频带上的特定时隙所对应的用户信道的例子，因此，这个系统中共有 NM 个信道。不久会看到北美 D-AMPS 或 IS-136 移动蜂窝系统中，一般每个用户对应一帧中的两个时隙。与每个用户占一个时隙相比，用户可能的比特率加倍了（每帧传输的比特数是原来的两倍），但相应地也减少了同时使用系统的用户数量。GSM 系统中每个用户对应每帧中的一个时隙，但是用户可以通过在不同频率信道中的移动来减小深度衰减的可能性（当然必须有控制策略来保证一个时间只有一个用户占有给定的信道）。

把 GSM 作为 FDMA/TDMA 系统的具体的例子，前面曾提到，在欧洲 GSM 系统被分配有两个 25 MHz 的频带。上行（反向，用户到基站）链路 890 MHz 到 915 MHz；下行（前向，基站到用户）链路 935 MHz 到 960 MHz（GSM 也可以适用于欧洲和北美的 2 GHz，这里只关注 900 MHz 系统）。每个 25 MHz 的频带依次分隔成 200 kHz 宽度的子频带，其中一个子频带用做保护频带，因此一个方向中实际上有 124 个可用频率。在第 5 章中提到的 GMSK 用做调制技术。每个 200 kHz 频带传送 8 分片的 TDMA 帧结构，帧以 4.615 ms 的间隔重复。如上所述，每个用户分配到每帧一个分片，所以，每个 GSM 系统中有 $124\times8=992$ 个可用信道。这个数量与图 6.2中的数量 NM 相对应（多帧结构也在系统中利用，这种结构将在第 8 章对 GSM 的详细介绍中描述）。

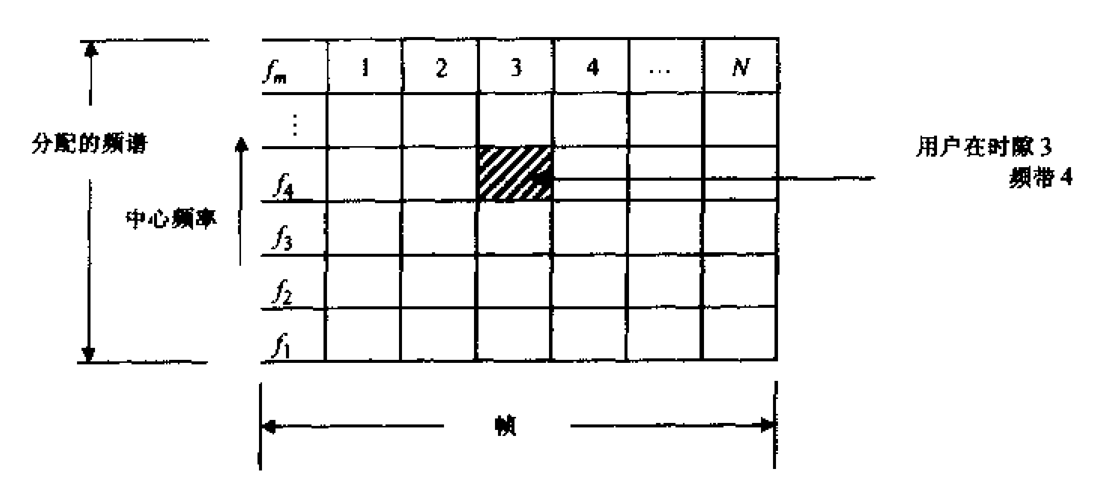


图 6.2 FDMA/TDMA 信道分配图

GSM 帧结构在图 6.3 中描述。可以看到每个时隙长度是 576.92 μs,包含 148 比特。时隙被长度为 30.46 μs 的保护时间分隔开,相当于 8.25 比特。每个时隙的 148 比特中,有 114 比特是数据比特,分成两个各有 57 比特的分组。3 个比特(称为 T 区域)定义了时隙的开始和结束。一个 26 比特的训练序列提供了必需的时间同步信息,还有两个 1 比特的标志位,这就是整个时隙的结构。系统的传输数据率是 154.25 比特/576.92 μs ≈ 270.833 Kbps。这个合成的比特传输对应于利用在 200 kHz 宽度频道上的 1.35 比特/Hz 的信道。回顾第 5 章曾指出 GSM 使用 0.3 GMSK 调制。0.3 是指 $B/R = 0.3$, B 代表低通带宽 3 dB, R 代表传输数据率。3 dB 传输带宽就是 $0.6R = 162.5$ kHz。200 kHz 的信道带宽是频带中点与之以下大约 10 dB 点之间的带宽(Steele,1992:540)[①]。

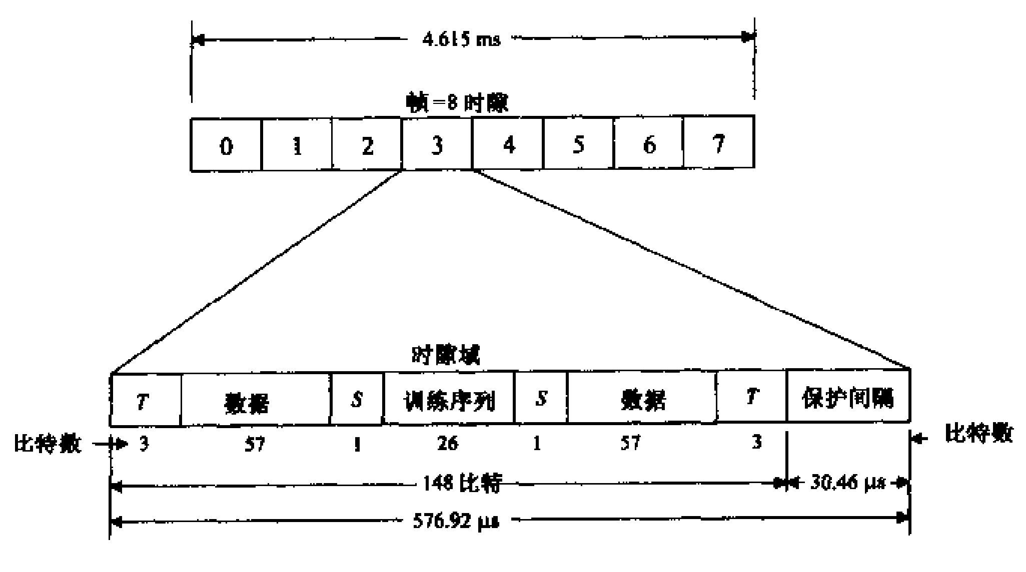


图 6.3 GSM 帧结构

用户数据率计算起来更复杂。每个用户要传输每帧 114 比特,然而,在第 8 章对 GSM 的详细讨论中会指出数据信道占有每 26 帧(复帧)的 24 个时隙。剩下的两个时隙用来传输控制信息。有效的用户比特率是 114/4.615 ms × 24/26 = 22.8 Kbps。这些比特是以270.833 Kbps的

① Steele, R. ed. 1992. *Mobile Radio Communications*, London, Pentech Press; New York, IEEE Press.

速率传输的。0.3 GMSK 调制方案加上跳频技术可使 GSM 达到 6.5 dB 的 SIR。这就使蜂窝 4 小区再用成为可能，而不是 IS-136 所要求的 7 小区再用。124 个可用频率的 1/4 被分配到每个小区，所以每个小区有 992/4 = 248 个可用信道。这个数目就是 GSM 的系统容量，并在后面本书讨论的重点内容——比较 3 种第二代系统中将用到。

IS-136 D-AMPS 蜂窝系统使用完全不同的 FDMA/TDMA 信道设置。本书已经指出，IS-136 使用与最初的模拟 AMPS 系统一样的频率分配方案：25 MHz 的频谱在每个传输方向被分成 832 个 30 kHz 的频带。但在 IS-136 中，第 5 章已经提到，使用的基本调制方案是 DQPSK。每个频带使用 TDMA 的方式重复传送每帧 6 个时隙的数据。图 6.4 表示了上行链路的帧时隙结构。帧长度是 40 ms，则以每秒 25 帧的速率重复。每帧的比特数是 1944 比特，则传输率就是 1944 × 25 = 48.6 Kbps。6 时隙的 TDMA 帧结构可以有两种用户：全速和半速用户。全速用户每帧分配两个时隙，因此每 30 kHz 的频率信道可以容纳 3 个这样的用户，或者 6 个半速用户，因为他们每帧分配一个时隙，所以传输速率是全速用户的一半。下面来看全速用户，他们可以每个时隙发送 260 个数据比特即每帧 520 个数据比特，从而提供的信息数据率是 520 比特/帧 × 25 帧/s = 13 Kbps，以 48.6 Kbps 的速率传输。对于语音传输，这个数据率可以有效处理以 7.95 Kbps 运行的语音编码器处理的编码压缩语音。正常的 64 Kbps 语音信号因此以大约 8:1 的比例压缩。第 8 章将介绍在这个系统中是如何处理语音（声音）信号的内容的。

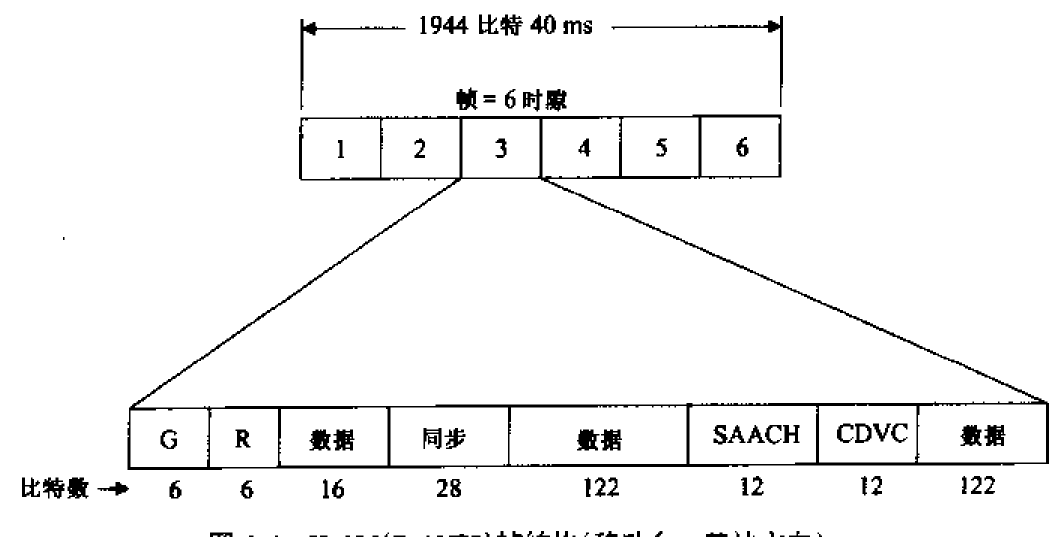


图 6.4　IS-136(D-AMPS)帧结构（移动台→基站方向）

图 6.4 中表示的时隙扩展的部分也会在第 8 章中做详细的介绍。在这里，只需要知道 G 是保护时间，R 为斜坡上升时间，同步区域是一个提供同步信息的特定的 28 比特序列，SAACH 区提供一个控制信道，而 CDVC 代表"编码的数字信道验证码"，是一个 8 比特数字加上 4 比特编码，编码用来使基站意识到有用户仍在活动：基站传输这个数字到移动台，移动台用同一个数字回应。如果移动台不发送这个数字，时隙即被放弃。

以上对 GSM 和 D-AMPS 的简单介绍包含了对时分多址接入 TDMA 的介绍。这一章剩余的部分会集中讨论码分多址接入技术 CDMA。在本章后面，当将 CDMA 系统与其他系统的用户容量进行比较时，会简要地回到 GSM 和 IS-136 系统。

6.2 码分多址接入技术

码分多址接入技术(CDMA)是蜂窝系统中用到的第3种多址接入技术,这种接入技术被用在北美的IS-95系统中。IS-95使用的是和AMPS以及IS-136(D-AMPS)相同的频谱分配。在第三代蜂窝系统中使用了更多的宽带CDMA版本。本书后面的第10章有更多关于这些先进系统的细节。CDMA是一种基于扩频技术的方案,这种技术多年前主要是为了军事通信系统而发明和发展的。

CDMA中用户之间的正交性是通过为每个用户分配不同的数字码来实现的。通过下面的方法来选择互相正交的码字:设一个码 i 由特定的 l 比特序列 $\{x_{ik}\}, k=1,\cdots,l$ 组成,二进制数位 $x_{ik}=\pm 1$。用元素向量 $\boldsymbol{c}_i$ 代表这个比特序列。正交性意味着点积 $\boldsymbol{c}_i \cdot \boldsymbol{c}_j = \sum_k x_{ik}x_{jk} = 0, j \neq i$。因此,多用户在分别被分配到不同码的情况下可以同时传输。码 i 的接收者通过执行点乘操作就可以惟一地恢复信号 i。只有在码 i 的时候才会出现输出,因为正交性,其他信号会被抑制。在实际中,我们会看到,用户代码是通过移位寄存器来产生的。产生的码的一个例子是伪随机序列(Viterbi,1995)[①]。图6.5表示了用这种伪随机序列产生CDMA系统的例子。图中表明使用了PSK的调制方式,也可以使用其他的调制方式(在第8章中将指出OQPSK和QPSK会在IS-95中使用)。

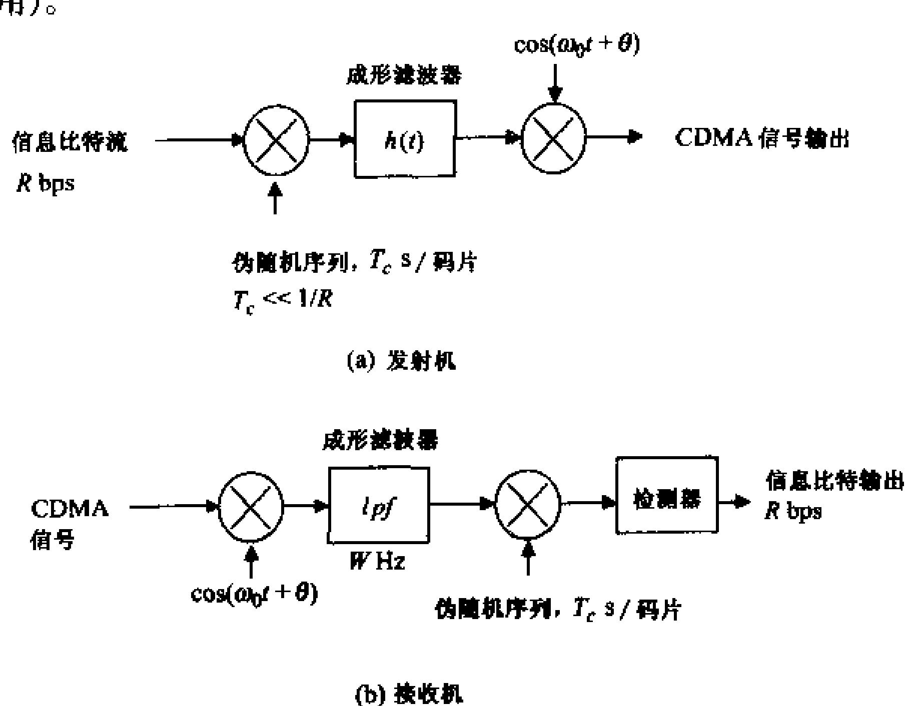


图6.5 CDMA系统发射机和接收机

伪随机序列如图6.5所示用来调制或乘信息比特流的每个比特。伪随机序列中的每个比特称为一片,长度为 T_c,所选择的长度比信息比特长度 $1/R$ 短很多,$T_c \ll 1/R$。伪随机序列乘以信息比特的目的是把二进制比特信息流转化成一个更宽频谱的类噪声序列,这也说明了扩频通信这个词可与缩写词CDMA交换使用。产生的宽带二进制序列的带宽 W 近似是 $1/T_c$。

① Viterbi, A. J. 1995. *CDMA, Principles of Spread Spectrum Communication*, Reading, MA, Addison-Wesley.

因此原始信息比特流的带宽要乘以 W/R，W/R 称为系统的扩频增益。马上可以看到，扩频增益越大，CDMA 系统的性能就越有效（实际中，伪随机序列中的二进制信号的运算是模 2 加法而不是乘法，两者是等同的。详见第 8 章）。

在接收端乘以相同的伪随机序列，如图 6.5(b)所示，重新产生原始的信息比特流。由于码元的正交性，其他使用相同载波频率的编码信号会被抵制。CDMA 的利用使全部频率的重用成为可能：相同的用户码可以在移动台漫游系统的任何单元中使用，这就简化了移动用户切换到新单元的信道的设定问题。但这又产生了另一个问题，由于所有单元的所有用户使用统一频率，增加了对一定基站或移动台的干扰（取决于传输方向）。除了增加的干扰功率，扩展的一个结果是提高了信号的可检测性，这会在本节后面论证。

伪随机这个词的意思是乘以信息比特流的二进制序列或用户码具有随机或类噪声特性，因此这样的序列常称为伪噪声(PN)。由于是由线性移位寄存器产生，所以该序列是确定性的。特别地，最大长度移位寄存器可能具有所要求的类噪声特性(Viterbi, 1995)。设移位寄存器的长度是 n，则它的第 m 个输出 a_m 可以由递归关系给出

$$a_m = \sum_{i=1}^{n} c_i a_{m-i}, \qquad c_i = 0, 1; c_n = 1 \tag{6.1}$$

这里的二进制输出是 1 或 0。0 必须被转化成 -1 来保证使用正交码。

最大长度移位寄存器定义成对所有非零原始向量而言间隔长度为 $P = 2^n - 1$ 的寄存器。随后会看出这样的移位寄存器具有如下类噪声特性(Pr[ ]代表[ ]的概率)：

1. $\Pr[0] = 1/2(1 - 1/P)$; $\Pr[1] = 1/2(1 + 1/P)$

   对于 $n \geqslant 10$，$1/P \leqslant 10^{-3}$，则 $\Pr[0] = \Pr[1] = 1/2$，因此 0 和 1 出现的概率大致是相等的，符合随机序列的期望特性。

2. 任何 l 比特长度的 0，1 序列的出现频率是 $1/2^l$，$l \leqslant n - 1$；$1/2^{n-1}$，$l = n$。例如，可以认为，在 0，1 随机序列中，有 1/2 是长度为 1 的序列；有 1/4 是长度为 2 的序列；有 1/8 是长度为 3 的序列。最大长度移位寄存器具有这种特性。

3. 对随机序列移位一定的元素数造成的与原来序列一致和不一致的数目相等，而最大长度移位寄存器对于因子 $1/P$ 也具有这样的特性(Viberbi, 1995)。这个特性意味着两个用户即使只移位一个码片，在 $1/P$ 范围内仍能够被单独检测。

最后的特性是用来规范 CDMA 正交码结构的有效性。具体来看用户 j 的第 k 个码片。前面已提到，它的值是 $x_{jk} = \pm 1$。设它与用户 i 的 l 个置换码片的值一致：$x_{jk} = x_{i,k+l}$。可以看到(Viterbi, 1995)

$$\frac{1}{P}\sum_{k=1}^{P} x_{ik} x_{i,k+l} = -\frac{1}{P}, l > 0; \quad = 1, l = 0 \tag{6.2}$$

图 6.5 中所表示的接收端的乘法器执行的就是这种相互关系。因此，在式(6.2)的基础上，信干比(SIR)的检测能力就是 $P = 2^n - 1$。这个值可以随着移位寄存器长度 n 按照足够大的要求来达到足够大。付出的代价是需要减小码片长度 T_c，把更多的伪随机比特挤进给定的 $1/R$ s 长度的信息比特中。传输带宽相应增加。式(6.2)所表示的 SIR 检测能力假设不同用户的码是校准的或同步的。然而，多移动台的上行链路传输却不是这么回事。缺乏码的校准以及因

此带来的码的正交性,会导致用户间干扰。在有许多干扰用户的条件下,这种干扰可以被认为是类噪声。在下面的章节中也使用了这个假设。

现在主要看一下 CDMA 可以达到的性能。主要是系统的容量,即给定系统可以容纳的用户数量。在 6.3 节中,通过只有一个单元干扰的情况来给出一个简单初步的 CDMA 系统容量的判定。同时,这对加强对 CDMA 系统的理解很有帮助。在 6.4 节中我们可以看到误比特率计算的概率结果,该结果是在研究 CDMA 系统容量所需的衰落环境中得到的。6.5 节又回到 CDMA 干扰的判定,考虑系统的所有用户的干扰。容量判定的最后结果用来比较基于 CDMA 的 IS-95 系统和基于 TDMA 的 GSM 与 IS-136 系统的容量。从下面几节的论述中,读者可以发现整个分析都是基于本书所使用的理想的系统几何模型、传输效应和衰落现象的。这些模型在提供对移动蜂窝系统实施环境的深入理解方面非常实用,也可以提供对这些系统所能达到的性能的一些度量。但是,在利用这些系统的结果来进行竞争系统相关性能的最后比较时需要谨慎。像这里所研究的移动蜂窝系统这样复杂的系统,在实际工作条件下可能与这里计算的结果差异很大。事实上,其结果随着地点、系统的执行以及其他因素面异。

6.3 CDMA 容量:单小区情况

上面曾提到,这一节来看一下怎么确定 CDMA 系统的容量,我们用最简单的单个小区中移动用户的干扰来说明。设小区中有 K 个用户,则每个用户有 $K-1$ 个干扰用户。我们主要看上行链路,从移动台到基站的方向,因为这是通信中最难控制的部分。假设小区中所有用户的功率都控制为与基站接收功率一样的功率 P_R。功率控制是 CDMA 系统性能的关键因素,否则,相邻的用户会有固有的优势。在第 4 章中讨论了功率控制技术的实例。

如果传输用户的数目足够大,每个都使用了自己的伪随机码,则可以假定合成的干扰接近高斯白噪声,从而基本的问题就变成了在噪声存在的情况下检测信号。这个问题已经研究了很多年。在通信理论中,我们都知道高斯白噪声(噪声的概率密度函数是高斯分布或正态分布且频谱为平坦的)下(二进制)信号的可检测性取决于接收信号的比特能量 E_b 与噪声功率谱密度 N_0(Schwartz, 1990①; Wozencroft 和 Jacobs, 1965②)的比。可检测性这个词指的是在一定错误概率下,能正确判别二进制信号(1 还是 0,或 1 还是 -1,根据具体情况而定)的能力。对于无线移动通信,噪声通常很小或可以忽略,其主导作用是类噪声的干扰。这决定于 $K-1$ 个干扰用户,干扰值是 $(K-1)\times P_R$,因此使用信号比特能量与干扰噪声功率谱密度的比来衡量信号的可检测性。

由于宽带编码 CDMA 信号的带宽值 W Hz $\gg 1/R$, R 是 6.2 节中提到的信息比特率,干扰的功率谱密度 I_0 用干扰功率除以频谱带宽来得到,即 $(K-1)P_R/W$。接收信号比特能量是 $E_b=P_R/R$。E_b/I_0 和信号比特能量与噪声密度之比 E_b/N_0 相当,决定了一般通信理论框架中的信号可检测性,公式如下:

$$\frac{E_b}{I_0}=\frac{P_R/R}{(K-1)P_R/W}=\frac{W/R}{(K-1)} \tag{6.3}$$

① Schwartz, M. 1990. *Information Transmission, Modulation, and Noise*, 4th edn, New York, McGraw-Hill.

② Wozencroft, J. M. and I. M. Jacobs. 1965. *Principles of Communication Engineering*, New York, Wiley.

扩频增益 W/R 越大，可容纳的用户 K 就越多。在可接受的误比特率基础上给定需要的 E_b/I_0 值，以及给定系统中的扩频增益 W/R，并忽略其他小区的干扰点，则一个小区中可以容纳的用户数量很容易得出，为

$$K = (W/R)/(E_b/I_0) + 1 \tag{6.4}$$

举个例子，设 $E_b/I_0 = 5$（或 7 dB），信息比特率 $R = 10$ Kbps，传输带宽是 $W = 1.25$ MHz（IS-95 系统数值）。在这个 1.25 MHz 的带宽中可以容纳 $K = 26$ 用户/小区。因此，北美蜂窝系统的 25 MHz 带宽的总用户数是 $20 \times 26 = 520$。可见，CDMA 系统中频率复用没有作用，因为每个用户都有各自的伪随机码，且所有小区总是可以使用所有频率；这也是通过使用扩频 CDMA 技术可以达到统一的频率复用的本质。

6.5 节将结合信号衰落效应、阴影衰落和所有系统用户带来的干扰等因素来改进这种简单的单小区干扰模型。正如可能期望的一样，我们将发现这种简单模型过于乐观了。在 6.5 节中更加综合的分析可以使我们能够比较 CDMA 和 TDMA 的系统容量，至少在使用的理想化模型的范围内。在讨论改进系统之前，首先看一下衡量信号可靠性的因子 E_b/I_0。在上面的实例中，为什么只选择 $E_b/I_0 = 5$？这个数字从何而来？在下一节（6.4 节）中，将总结一些来自通信理论中关于误比特率性能的数字，这些理论比较了不同调制方案中在高斯噪声下的二进制信号的可靠性。我们也会举出一些数字把性能评价扩展到第 2 章中所讨论的那种瑞利快速衰落信道上。对于之前具备通信理论背景的读者来说，这些结果可能成为他们前面工作的一个回顾。对于之前未具备通信理论背景的读者，就要单凭相信来接受这些结果，或找到分析中使用的参考资料来查看相关的细节了。

6.4　误比特率的考虑

在这一节的开始先通过归纳通信理论中各种常见调制方式下信号可靠性的一些相关结果，来看一下误比特率以及它和移动系统二进制信号可靠性的关系。

例如，可以看出 PSK 是在使用最简单通信信道情况下的性能最好的二进制调制方式。在这种调制方式中加性高斯白噪声，即假定在传输过程中加到信号上的高斯统计特性的噪声，是信号可靠性的惟一威胁因素（Schwartz，1990；Wozencroft 和 Jacobs，1965）。性能这个词是指可以达到的误比特率，即二进制数错误接收的比率。性能最好的方案就是可以达到最小错误率的方案。在 PSK 中，在加性高斯白噪声存在的情况下的误比特率 P_e，根据互补错误函数容易得出，公式如下（Schwartz，1990：475）：

$$P_e = \frac{1}{2}\mathrm{erfc}\sqrt{\frac{E_b}{N_0}} \tag{6.5}$$

这里 E_b 是接收信号的比特能量；N_0 是噪声功率谱密度，它们都在前面定义过。互补错误函数 erfc (x) 通过下面的积分式定义：

$$\mathrm{erfc}(x) \equiv \frac{2}{\sqrt{\pi}}\int_x^{\infty} \mathrm{e}^{-y^2}\mathrm{d}y \approx \frac{\mathrm{e}^{-x^2}}{x\sqrt{\pi}}, \quad x > 3 \tag{6.6}$$

由式（6.5）和式（6.6），可以得到如下对小概率值 PSK 误比特率的表达式：

$$P_e \approx \frac{1}{2}\frac{\mathrm{e}^{-E_b/N_0}}{\sqrt{\pi E_b/N_0}} \tag{6.7}$$

可见,错误率随着指数 E_b/N_0 反向变化。比如,设要达到的 $P_e=10^{-5}$,也就是平均传输 10^5 个数中有一个接收错误。由式(6.7)可以得出 E_b/N_0 用分贝单位计算是9.6 dB。将 E_b/N_0 减小一半,或减小3 dB为6.6 dB,则可得 $P_e=10^{-3}$。由此可见,由于指数 E_b/N_0 的变化,P_e 有了两个数量级的增加!因此,错误率性能主要取决于接收信号比特能量和噪声频谱密度的比值 E_b/N_0。

如果使用FSK调制方式,也使用相同的加性高斯白噪声,会发现出现了一个3 dB的损失:所需的 E_b/N_0 是原来的两倍,或增加了3 dB。对于固定噪声谱密度,这意味着要达到与PSK相同的误比特率,所需的信号比特能量或功率必须加倍。比如,如果需要 $P_e=10^{-5}$,用FSK调制,$E_b/N_0=12.6$ dB;对于 $P_e=10^{-3}$,$E_b/N_0=9.6$ dB。在第5章中曾提到,PSK和FSK需要发送端和接收端之间准确的相位同步。但我们也可以在比较两个传输频率的相对信号幅度和包络的基础上,通过非相干检测FSK信号,而不必使用相位同步(这对PSK明显是不可能的)。比如,需要 $P_e=10^{-5}$,则其代价是0.7 dB的损失,也就是非相干FSK需要增加0.7 dB的信号能量或功率——所需的 E_b/N_0 增加到13.3 dB。我们也可以看到差分PSK或DPSK需要比PSK多一个分贝的信号功率:对于 $P_e=10^{-5}$,$E_b/N_0=10.5$ dB(这种调制方案与第5章讨论的DQPSK类似)。在载波调制之前进行信号编码可以显著地改善这些数字。比如,如果使用编码率是1/2的卷积编码,调制方式是PSK,则此时需要的信号能量与噪声密度之比将从4 dB变为6 dB。根据编码方式的不同,在上面的9.6 dB基础上会有明显的减小(Viterbi, 1995; Schwartz, 1990:693)。编码技术(包括卷积编码)将在第7章中讨论。

现在考虑衰落环境下的情况。这些数字会明显地变化(Schwartz 等,1966)[①]。PSK的误比特率大概是

$$P_e=1/4(E_b/N_0) \tag{6.8}$$

这个反比例关系与非衰落环境中式(6.7)的反指数关系形成对比。这意味着,即便 $E_b/N_0=10$ dB,误比特率大概是1/40或0.025,也比前面的 10^{-5} 增加了3个数量级。DPSK和FSK都需要比PSK多3 dB的信号功率,与非衰落环境中的性能减少量一致。第2章中讨论的分集方案可用来补偿由于衰落(Schwartz 等,1966)造成的比较严重的性能退化。例如,调制方式为FSK,使用双重分集,把两个接收天线的输出和每个天线的接收信号合并。假定每个独立衰落,发现这样能够以一定的量级阶数提高误比特率性能。对于 $E_b/N_0=10$ dB,P_e 从非分集模式的0.05下降到双重分集的0.005[忽略合并方案的类型。在第2章中曾讲到,可以使用选择分集选择多重接收有效信号中的最大信号;简单地把接收多重信号加在一起,也就是说,等增益合并;或者使用最大比合并,将每个接收信号用其信噪比加权。在第2章中也曾讲到,这种方案可能提供最好的分集性能(Schwartz 等,1966)]。

让我们回顾一下,这些用来提高衰落环境中系统性能的分集技术已经被人们认识了许多年,这也尽可能地提前了它们在移动无线系统中的使用。使用时一般将两个或多个接收天线联合使用,就像在上面的例子中那样。然而,最近有研究指出,通过在发送端和接收端同时使用多天线可以使传输容量得到明显而实际的提高。现代数字信号处理(DSP)技术使多天线具

① Schwartz, M., W. R. Bennett, and S. Stein. 1966. *Communication Systems and Techniques*, New York, McGraw-Hill; reprinted, IEEE Press, 1996.

有了实用性。结合多天线技术的系统称为多输入多输出（MIMO）系统（Matsumoto 等，2003）[①]。MIMO 方案可以与最近在编码技术上的一些进步如 turbo 编码相结合，turbo 码会在第 7 章中讨论，其合成系统使用空时编码。这些最近的发展使人们对它们在实际无线系统中的应用产生了很大的兴趣。空时编码已经被提议用在第 10 章所讨论的第三代 CDMA 移动通信系统中。一个称为 BLAST 的前期 MIMO 执行方案已经在贝尔实验室研发出来，用在 CDMA 系统中（Foschini，1996）[②]。仿真研究表明，BLAST 可以潜在地提供传输容量上的非常高的增加量。

现在具体看一下扩频通信的情况，本章中的 CDMA 系统就是例子。回顾一下，第 2 章中描述了一个分集方案的特殊例子称为瑞克接收机，它利用了一个事实，即对带宽非常宽的信号，同一信号的不同版本可能在不同的时间到达，还可以被分别解决。瑞克接收机把这些单独的多径信号合并，结果是实际的分集改进（Price 和 Green，1958）[③]。但是事实上 CDMA 由于伪随机码确实产生了非常宽的宽带信号，以及高扩频增益。换句话说，用码片值比数字符号间隔小得多的伪随机编码，使信号变换成为比数字符号间隔和码片大得多的延迟扩展。计算表明，用瑞克接收机进行卷积编码的 CDMA 信号，其双重分集接收需要 E_b/I_0 的比例为 5(7 dB)，以达到 $P_e=10^{-3}$ 的误比特率（Gilhausen 等，1991[④]；Viterbi，1995）。这个数字表明对于双重分集 FSK，该数字比前面得到的数字有明显的改进（注意，回到使用干扰频谱密度 I_0 而不是噪声频谱密度 N_0）。如果 E_b/I_0 减小 1 dB 到6 dB或者数值为 4，则误比特率 P_e 会增加一个数量级到 0.01！可见信号能量与干扰频谱宽度的比值是系统性能的决定性因素。

在表 6.1 中总结了瑞利多径衰落对误比特率的影响和利用分集技术来减轻这种效应的一些评价。在表中，适用于本章所讨论的 CDMA 方案的数值用粗体表示。总而言之，这个关于信号比特能量与噪声或干扰频谱密度的比值，以及其与误比特率相互影响的简单讨论表明了 CDMA 系统的容量，即每单元允许的用户数量。它既取决于认为可接受的误比特率的选择，也取决于为了克服衰落效应而采用的编码/分集方式。在本章剩下的部分将一直使用 $E_b/I_0=5$ (7 dB)这个值。其他的作者使用了 $E_b/I_0=4$(6 dB)的值，这个值在上面曾提过，其对应的误比特率是 10^{-2}。注意，第二代蜂窝系统最初设计用来且现在也主要用来进行声音传输。数字语音对误码率要求相对鲁棒一些。这就是为什么误码率 $P_e=10^{-2}$ 或 10^{-3} 也可以提供可接受的性能，而这些数字对于数据传输是不能接受的。

表 6.1　E_b/I_0 vs 错误比特概率；在不同的通信信道下

调制类型和信道	需要的 E_b/I_0(dB)				
	$P_e\rightarrow$	0.01	0.005	0.001	10^{-5}
PSK					
无衰落，无编码		4		6.6	9.6
卷积编码					4～6

① Matsumoto, T. *et al*. 2003. "Overview and recent challenges of MIMO systems," *IEEE Vehicular Technology News*, 50, 2 (May), 4－9.

② Foschini, G. J. 1996. "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Laboratories Technology Journal*, 1, 2, 41－59.

③ Price, R. and P. E. Green, Jr. 1958. "A communication technique for multipath channels," *Proc. IRE*, 46, 555－570.

④ Gilhausen, K. H. *et al*. 1991. "On the capacity of a cellular CDMA system," *IEEE Transactions on Vehicular Technology*, 40, 2 (May), 303－312.

（续表）

调制类型和信道	需要的 E_b/I_0(dB)				
	$P_e \rightarrow$	0.01	0.005	0.001	10^{-5}
衰落信道		14	17	24	
使用双分集 RAKE 和卷积码		6		7	
DPSK					
无衰落					10.5
衰落信道		17	20	27	
FSK					
无衰落		7		9.6	12.6
衰落信道		17	20	27	
使用双分集			10		

6.5 CDMA 容量计算:CDMA 与 TDMA 相比

6.3 节中讨论了 CDMA 系统中每个单元可以容纳的用户数量,使用的是一个简单的单小区模型,并且不包含衰退效应和阴影衰落。在这一节中推广一下包含这些效应的计算。具体地,包含了不仅是单个小区,而且是遍及系统区域互相干扰的用户的效应,还有阴影衰落的效应,使用第 2 章中所描述的模型(快速多径衰落假定同双重分集瑞克接收机一同考虑,包含在 E_b/I_0 的值的选择中,使用 6.4 节中讨论的数字)。主要关注上行链路,从移动台到基站方向。在 6.3 节中曾提到,CDMA 系统的操作主要取决于功率控制。假设所有移动台的功率被与它们通信的下行链路信号所控制,因此,任何基站接收的上行链路信号都有相同的平均功率值 P_R。回顾第 4 章中描述的功率控制技术,就是用来设计符合这种条件的。同时,假设阴影衰落变化的足够缓慢,则基站功率控制可以处理接收功率由于这种衰落而引起的变化。

首先,主要看的是硬切换:在一个给定小区中一个移动用户只与一个基站通信,但它穿过小区边界进入另一个小区就能够立即切换,或者移交到新的基站。移交策略的分析在第 9 章中将详述。然而,CDMA 也允许软切换。在这种情况下,一个移动用户可能与两个或多个基站通信。用户将会被分配给衰退损失最小的那个基站,这样可以减小整体干扰功率,提高系统容量,即每个小区中允许的同时使用的用户数量。用来分析的小区模型假设是六边形小区(后面可以看到,基于圆形小区结构的分析得到的是本质上相同的结果,这意味着结果在大体上对假设的小区几何形状具有鲁棒性)。假设一个均匀的条件:用户在小区中均衡分布,各处的用户密度一致;遍及系统中所有的接收功率有相同的数值 P_R;整个系统每个小区的衰退损失和阴影衰落系数相同。这是一个非常理想化的模型。

看一下图 6.6 中六边形小区的结构。采用 Viterbi(1995)的方法来计算基站的整个干扰功率。由于系统定义成均衡的,可以选取一个典型的小区 S_0,其基站位于中心位置点 0 处。对于在 6.3 节中曾经被忽略的位于这个小区之外的干扰移动用户,也就是在 S_0 之外的所有小区里的用户,我们把它们所在的区域记为区域 S_0^*。这些移动用户,在向他们自己的基站 BS_0 传输信号的过程中,也对基站产生了干扰。需要考虑的是,外部区域 S_0^* 的整个干扰功率。举个例子,假设小区 1 中一个位于点(x,y)的移动用户,如图所示,它与其基站 BS_1 进行通信,但在

基站 BS_0 也接收到该用户的干扰功率。上面说到，假设用户在一个小区中均匀分布，就是说，有一个均衡的通信量密度 ρ 用户/m²。由于半径为 R 的六边形小区的面积是$(3\sqrt{3}/2)R^2$（见第 3 章），假设每小区有 K 用户，则通信量密度为

$$\rho = 2K/3\sqrt{3}R^2 \tag{6.9}$$

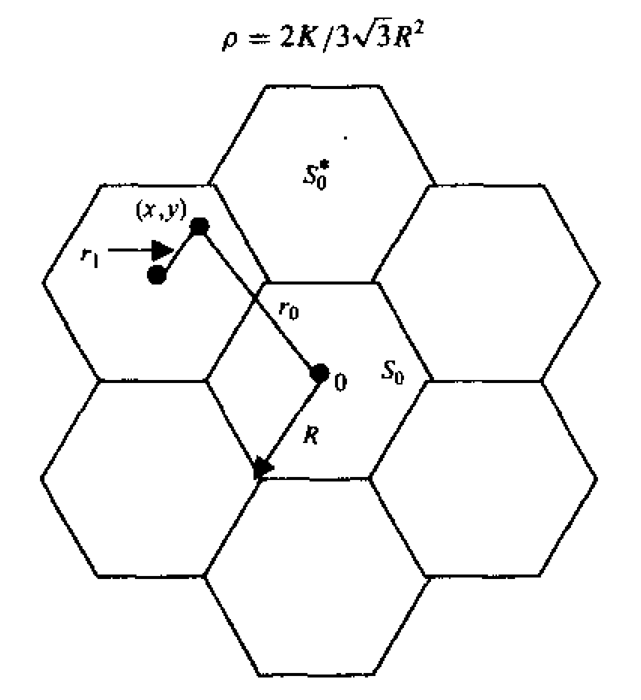


图 6.6　CDMA 外部干扰功率的计算

假设系统中每个基站需要相同的接收功率 P_R，这决定了移动台到基站传输 r m 距离需要的功率。回顾第 2 章式(2.4)，忽略了快速衰落的接收功率如下：

$$P_R = P_T r^{-n} 10^{z/10} \tag{6.10}$$

P_T 是移动台的传输功率；n 是衰退指数，通常值为 4；z 是高斯分布阴影衰落的随机变量，以平均功率为中心，标准差为 σ，单位为 dB。其概率密度函数为

$$f(z) = \mathrm{e}^{-z^2/2\sigma^2}/\sqrt{2\pi\sigma^2} \tag{6.11}$$

现在主要看点(x,y)的干扰移动台，距小区 1 中基站的距离是 r_1。设阴影衰落的随机变量为 z_1，则移动台的传输功率为

$$P_{T_1} = P_R r_1^n 10^{-z_1/10} \tag{6.12}$$

（注意，由于阴影衰落随机变量 z 关于 0 对称分布，在式(6.10)中写成 $10^{-z/10}$ 和这里的 $10^{z/10}$ 是等同的）。那么，在此基站 BS_0 从这个距基站 BS_0 为 r_0 的移动台接收到的（干扰）功率可以表示如下：

$$P_{T_1} r_0^{-n} 10^{z_0/10} = P_R \left(\frac{r_1}{r_0}\right)^n 10^{(z_0-z_1)/10}$$

在一个不同区域 $\mathrm{d}A(x,y)$，以点(x,y)为中心，移动台的数目为 $\rho \mathrm{d}A(x,y) = 2K\mathrm{d}A(x,y)/3\sqrt{3}R^2$。用这个数来代替前面单个的干扰用户，综合整个 S_0^*，取随机变量的平均值或期望值（用符号 E 表示），则在基站 BS_0 的外部区域 S_0，其移动台引起的全部平均干扰功率可由下式得出（Viterbi，1995）：

$$
\begin{aligned}
I_{S_0^*} &= \frac{2K}{3\sqrt{3}R^2}P_R E\left\{\iint\limits_{S_0^*}\left[\left(\frac{r_1}{r_0}\right)^n 10^{(z_1-z_0)/10}\right]\mathrm{d}A(x,y)\right\} \\
&= \frac{2K}{3\sqrt{3}R^2}P_R E\left[10^{(z_0-z_1)/10}\right]\iint\limits_{S_0^*}\left(\frac{r_1(x,y)}{r_0(x,y)}\right)^n \mathrm{d}A(x,y)
\end{aligned}
\tag{6.13}
$$

则外部干扰功率的表达式可看成两部分的乘积：一部分是单纯的几何部分，包含外部区域 S_0^* 的综合；另一部分是阴影衰落的平均。单纯的几何二重积分的数字积分值，在计算中包含系数 $2/3\sqrt{3}R^2$，当 $n=4$ 时是 0.44(Viterbi，1995)（后面显示了使用圆形小区的积分结果是一个非常相似的值）。现在来看阴影衰落部分，这一项非常重要，它为这里使用的模型提供了一个相对更大的值。因此，阴影衰落显著地增加了干扰功率（几何项的效果对衰退参数 $n=4$ 的情况相对消失得很快，所以小区 S_0 周围第 1 层的小区占了干扰功率的大部分。将直接使用圆形模型来说明这种效果）。

现在看阴影衰落表达式(6.13)的计算问题。对这个问题的分析第 1 次在 Viterbi 等人于 1994 年出版的著作①中进行了描述，在 1995 年的版本中也同样出现过。两个阴影衰落随机变量 z_0 和 z_1 代表两个基站 BS_0 和 BS_1 所测量的各自功率的变化额。注意，由于这些基站的接收功率是由于点(x,y)附近的移动传输引起的，因此这个区域的阴影衰落效应必须也考虑在内。在 Viterbi 等人于 1994 年和 1995 年出版的著作中，假设每个基站所测量的阴影衰落随机变量(rv)通过两个随机变量的和给出。一个随机变量对两个阴影衰落项都适用，代表在点(x,y)附近传输的移动台的阴影衰落效应。第 2 个随机变量代表在衰退路径中的随机功率变化，并设定在从(x,y)到 BS_0 和 BS_1 的两条路径中是独立的。具体而言，这个假设把两个阴影衰落随机变量 z_0 和 z_1 写成下面的形式：

$$z_i = ah + bh_i \qquad i=0,1 \quad a^2+b^2=1 \tag{6.14}$$

随机变量 h 代表传输移动台附近的阴影衰落，因此对问题中的两个基站的接收功率是共同的，基站 BS_1 对另一个基站 BS_0 来讲是干扰因素。另一个随机变量 h_i，$i=0$ 或 1，代表由于两条路径的衰退条件而加上的阴影衰退。如果阴影衰落的一半是由于传输移动台的共同区域造成的，则另一半是由于两个基站的阴影衰落的独立性造成的，满足 $a^2=b^2=1/2$。注意，这里所讨论的在式(6.14)中的 3 个不同的随机变量 h，h_0 和 h_1，必须各自满足高斯分布且独立。因此，第 1 个和第 2 个评估如下：

$$
\begin{gathered}
E(z_i)=0=E(h)=E(h_i);\quad E(z_i^2)=\sigma^2=E(h^2)=E(h_i^2); \\
E(hh_i)=0=E(h_0h_1)
\end{gathered}
$$

现在考虑式(6.13)中阴影衰落表达式的指数的计算。首先，在式(6.14)中，$(z_0-z_1)=b(h_0-h_1)$是高斯分布的随机变量，均值是 0，方差是 $2b^2\sigma^2$。为了简化式(6.13)中指数的计算，定义变量转换式

$$e^y \equiv 10^{(z_0-z_1)/10} \tag{6.15}$$

① Viterbi, A.J. *et al*. 1994. "Soft handoff extends CDMA cell coverage and increases reverse link coverage," *IEEE Journal on Selected Areas in Communications*, 12,8 (October), 1281 - 1288.

其中,$y=(z_0-z_1)/10\times\ln10=0.23(z_0-z_1)$,也是高斯分布,均值为 0,方差为 $\sigma_y^2=(0.23)^2\times 2b^2\sigma^2=0.053\sigma^2$,若根据上述建议,取 $b^2=1/2$。现在,式(6.13)中的指数项就很容易计算出来了,如下所示:

$$E\left[10^{(z_0-z_1)/10}\right]=E(e^y)=\int_{-\infty}^{\infty}e^y\frac{e^{-y^2/2\sigma_y^2}}{\sqrt{2\pi\sigma_y^2}}dy=e^{\sigma_y^2/2} \tag{6.16}$$

举个例子,在普通的阴影衰落情况中取 $\sigma=8$ dB[Viterbi(1995)列出了其他情况的数据表,包括改变衰退指数 n 的效果。为了保证讨论的简明,只关注这一个例子中的情况]。在该例子中,$\sigma=64$,$E(e^y)\approx5.42$。注意,这个值很大,在前面式(6.13)随后的讨论中提到过。可见,阴影衰落对系统容量有很大的影响。用这个值作为阴影衰落的标准差,以及式(6.13)中的双重积分的值 0.44,最后可以得到外部小区干扰的平均功率

$$I_{\overline{S_0}}=P_RK\times0.44\times5.42\approx2.38P_RK \tag{6.17}$$

这个代表一个小区外部的干扰功率要加到小区内干扰功率$(K-1)P_R$上,这来自于 6.3 节中讨论的小区内部的 $K-1$ 个移动台的干扰。外部干扰功率比内部干扰功率大,主要是由于外部的阴影衰落效应。对于普遍使用的 $n=4$,$\sigma=8$ dB 的情况,两项干扰的和即全部干扰功率为 $I=P_R(3.38K-1)$,而对于 W-Hz 的传输带宽,对应的干扰频谱密度为 $I_0=(3.38K-1)P_R/W$。信号比特能量与干扰频谱密度之比 E_b/I_0,在这里对应的值变成

$$E_b/I_0=(W/R)/(3.38K-1) \tag{6.18}$$

将这里 E_b/I_0 的值跟 6.3 节中的值进行比较。如果小区外部干扰和阴影衰落使小区中的用户数减少超过 3 个,就会使 E_b/I_0 下降,或者反过来说,要保持 E_b/I_0 一定的值,就要把用户的数目减少超过 3 个。具体来说,取之前选择的数值 $E_b/I_0=5$(7 dB),令 W 为 IS-95 的传输带宽 1.25 MHz,信息比特率还是 6.3 节中例子的数值 $R=10$ Kbps,可得到每个小区用户数 $K=7$ 或 8。这与在先前的计算中只考虑小区内干扰的每小区 26 个用户的数量相比,在容量上有一个明显的减少(对于全部的 25 MHz 带宽,这意味着每小区可以容纳平均不超过 160 个用户)。不过,我们忽略了使用 CDMA 系统可以带来的一些系统容量上的改进。这些因素包括前面提到的软切换,小区扇形分区的可能性(3 个天线每个覆盖 120°的扇区,理想情况下可把干扰者数量减少至 3 个)和语音静默检测。这些与它们对系统容量计算的影响后面会讲到。但是,首先停下来看一下式(6.13)中对六边形小区得到的双重积分数值 0.44。前面说过,圆形小区模型可以得到几乎一致的结果 0.43。这至少证实了对小区几何形状计算的鲁棒性。这里采用的是 Kwok 和 Wang(1995)[①]提出的分析结果。这篇论文针对六边形小区给出了对小区外干扰的独立计算,忽略了阴影衰落,得到了与 Viterbi(1995)相一致的结果。

看一下图 6.7 中的圆形干扰模型。这个圆形模型与图 6.6 的六边形干扰模型相当。图 6.6 表示了小区 S_0 周围第 1 层的 6 个干扰小区中的一个。为了使圆形模型与六边形模型相当,使半径为 R_S 的圆形面积与半径为 R 的六边形面积相等。可列式 $\pi R_S^2=3\sqrt{3}R^2/2$,得到 $R_S=0.91R$。从图 6.6 中还可以看出,两个小区的中心点,点 0 和点 1 之间的距离是$\sqrt{3}R$。从图 6.7

① Kwok, M.-S. and H.-S. Wang. 1995. "Adjacent-cell interference analysis of reverse link in CDMA cellular radio system," IEEE PIMRC95, Toronto, Canada, September, 446 - 450.

中也可以看出相同的距离。为了表示圆形模型中的计算方法,我们只进行第1层干扰小区引起的"小区外部干扰"的计算。有意思的是这个分析很容易得出,得到的是一个封闭的表达式,而不用像计算六边形模型那样需要计算式(6.13)中的双重积分。计算图6.7中表示的第1层圆形小区散发的总的干扰噪声功率,然后由对称性乘以6得到总的第1层的干扰功率。圆形使这个计算相对简单得多。用式(6.13)表达相同的方法,假定图6.7的圆形小区中 K 个干扰移动台在小区面积 πR_S^2 中均匀分布,我们还是使用 r^{-n} 的衰落定律,并用功率控制,则任何基站接收到的移动台的功率具有相同的值 P_R。先计算在点0的基站接收到的微分区域 dA 中,点(x,y)附近移动台的干扰功率,但现在使用极坐标。使用图6.7中的极坐标,$dA = r_1 dr_1 d\theta$。综合整个干扰小区并乘以6,得到整个第1层的干扰功率。可以看到第1层干扰的表达式与式(6.13)类似,且忽略了阴影衰落,具体如下:

$$I_{1st,S_0^*} = \frac{6KP_R}{\pi R_S^2}\int_0^{R_S}\int_0^{2\pi}\left(\frac{r_1}{r_0}\right)^n r_1\,dr_1\,d\theta \tag{6.19}$$

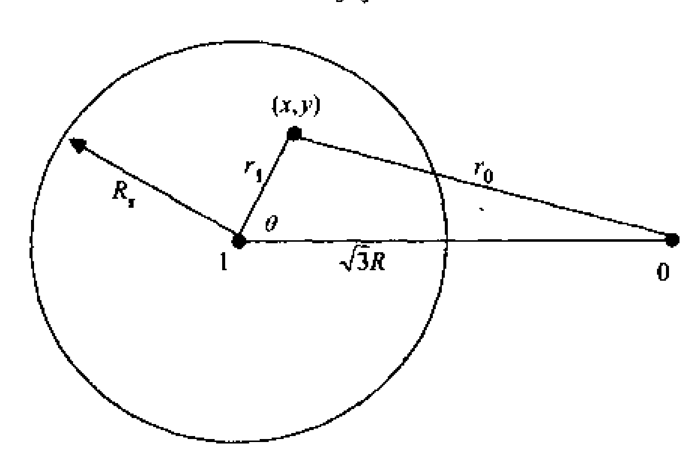


图6.7 圆形小区模型外界干扰功率计算

现在,在式(6.19)中,从微分区域中的点(x,y)附近的干扰移动台到小区 S_0 中的点0的基站之间的距离变量为 r_0。可以根据图6.7的结构,用极坐标 r_1 和 θ 的形式写成

$$\begin{aligned} r_0^2 &= \left(\sqrt{3}R - r_1\cos\theta\right)^2 + r_1^2\sin^2\theta \\ &= \left(3R^2 + r_1^2\right) - 2\sqrt{3}r_1 R\cos\theta \end{aligned} \tag{6.20}$$

用式(6.20)替代式(6.19)中的 r_0,首先对 θ 积分,然后将剩下的对 r_1 的积分式用虚拟变量 $x = r_1/\sqrt{3}R$ 代替 r_1。对于 $n=4$ 的情况,可以得到

$$I_{1st,S_0^*} = 36KP_R\int_0^{0.525}\frac{x^5(1+x^2)}{(1-x^2)^3}dx \tag{6.19a}$$

读者可以看到,式(6.19a)中的积分很容易计算,可得到第1层干扰功率

$$I_{1st,S_0^*} = 0.38KP_R \tag{6.19b}$$

这个值用来与六边形中对式(6.13)的双重积分在整个空间上得到的值 $0.44KP_R$ 相比较。用六边形来计算,等同的第1层数值为0.384,这在Kwok和Wang(1995)的著作中计算得到。对于

其他层，前面已提过，Kwok 和 Wang 得到的数值为 $0.43KP_R$。因此，圆形和六边形模型的结果几乎相同，又一次像前面那样表明这个计算对几何形状具有鲁棒性。也可以注意到第 1 层的结果与整个空间上的积分结果很接近，这显然是由于假设 r^{-4} 衰退关系式而引起的。远处的小区干扰效应随着距离的增加迅速衰减为零。在前面章节计算 SIR 时也有这样的结果。

现在回到 CDMA 系统所能承载用户数的计算问题上，对于一个给定的信号比特能量与干扰频谱密度的比值 E_b/I_0，计算其用户数。前面已经表明，如何计算外部小区干扰，包含阴影衰落效应。现在来讨论考虑使用如软切换、语音静默检测和天线扇区等附加效应而可能实现的 CDMA 系统容量的改进。在上面的讨论中建立了假定是硬切换的模型的干扰的计算。由于 CDMA 需要功率控制，前面讨论的软切换成为可能，则移动台在两个或多个基站中需要决定哪个用来与之进行通信。计算表明，如果移动台选择两个基站中较好的一个，则外部小区干扰项可从式(6.17)中的 2.38 减小到 0.77(Viterbi, 1995)。如果移动台可以选择 3 个基站中最好的一个，则该项可进一步减小到 0.57(Viterbi, 1995)。这带来了用户数量的一个可能的增加 3.38/1.77，或者大约二比一[见式(6.18)]。这个结果的假设条件是理想的功率控制。1 dB 的功率控制损耗系数是典型的，但却把可能的改进减小了 1.25(Viterbi, 1995)。使容量明显地进一步增加成为可能的是自动语音静默检测：众所周知，正常说话时，说话人在非常短的无声间隔与活动话音产生或谈话迸发之间进行轮流交替(Schwartz, 1996)①。一个谈话迸发持续的典型时间从0.4 s 到1.2 s，然后跟着一个从 0.6 s 到 1.8 s 的无声间隔。平均说话时间占整个时间的 0.4。语音电话网络几年来利用这个现象允许大约是原来正常数量 2.5 倍的用户进入网络。在传输语音信号的 CDMA 系统中，无声意味着没有功率传输，相应地自动减少了干扰功率。对于语音电话的情况，不需要语音静默检测设备。因此，由于语音静默检测可以减小 CDMA 干扰功率大约为 2.5[Viterbi(1995)使用的是 2.67]。最后，假设基站使用 120°的扇形分区。即使重用的系数是 7，AMPS 和 D-AMPS 系统要求使用天线分区。在 CDMA 中，天线分区的使用使干扰功率可能减小 3，因为每个天线的干扰可以只来自相应扇区的小区外部移动台。但这个值假定是在理想的基站天线的情况下。允许 1 dB 的天线增益损失会使可能的改进由 3 除以系数 1.25 减小到 2.4。

把上面描述的不同因素放在一起，传输带宽为 1.25 MHz 的 IS-95 CDMA 系统容量的最终结果是如果使用两小区软切换，则每小区。容纳 84 个移动台；如果使用 3 小区软切换，则每小区可容纳 96 个移动台。这些结果是针对阴影衰落模型的情况得到的。在该模型中，$\sigma = 8$ dB，衰退指数 $n = 4$，若要求 E_b/I_0 为7 dB，则相应的值是 5。如果 E_b/I_0 可减小到 4(6 dB)，也就是说，相应的误比特率的增加是可以容忍的，则相应的每小区的容量值可以增加到 108 和 120。回顾最初的假定：这些结果是用蜂窝系统的理想化模型得到的，并且这些都是平均容量值。一些统计性能的衡量在第 2 章和第 4 章中进行了简要的描述，更多的统计性能的衡量细节出现在第 9 章的性能分析中。

对其他一些不同于这里选择的参数值，如衰退常数 n 和阴影衰落标准差，以及软切换效应的 CDMA 系统容量的计算出现在 Vierbi 的著作(1995)中。这里讨论的 CDMA 系统结果在表 6.2中进行了总结，并与相应的 1.25 MHz 带宽的 AMPS，D-AMPS 或 IS-136 以及 GSM 系统的容量值进行了比较。AMPS 和 D-AMPS 得到的结果是假定重用率为 7；GSM 得到的结果假定重

① Schwartz, M. 1996. *Broadband Integrated Networks*, Englewood Cliffs, NJ, Prentice-Hall.

用率为4。读者可以自己讨论设定相关细节(注意,在前面6.1节对于TDMA系统容量的讨论中,使用的是全部25 MHz带宽的容量。因此,得到的数据必须减小为原来的1/20,以得到表6.2中1.25 MHz频带范围的数据)。可以看到,在这里的理想化模型的基础上,CDMA系统对比TDMA系统提供了充实的平均容量的改进。CDMA性能对比TDMA系统计算出的性能的改进已经被争论了多年。几乎所有的容量改进都是由于软切换结合自动语音静默检测的使用带来的,语音静默检测也可以应用到TDMA系统中,但代价是系统的复杂性增大。对于结合了语音静默检测的TDMA系统的仿真,研究表明可以提供与CDMA系统类似的系统容量性能(Gudmundson等,1992[①];Sköld等,1995[②])。尽管对于CDMA计算的争论假设使用的是理想模型,但本作者的观点是这仍是一种很有意义的活动,因为它确实注重CDMA系统的实施以及系统引进的各种特性。可以在第10章中看到,CDMA已经是第三代传输高速数据或语音蜂窝系统阶倾向使用的一种多址接入技术。主要原因是它能够提供灵活操纵两种类型信息的能力。扩频技术也被用在本地领域、个人领域以及ad hoc无线网络中,这些在第12章将会讨论到。

表6.2 在移动终端、小区,1.25 MHz带宽情况下容量的比较

AMPS[1] (analog)	D-AMPS[1] (TDMA)	GSM[2] (TDMA)	IS95[3] (CDMA)
6	18	12~13	84~96

注意:1. 重用因子为7,120°扇区天线。

2. 重用因子为4。

3. $E_b/I_0=7$ dB,$n=4$,$\sigma=8$ dB,语音静默检测因子为2.4,2或3小区软切换,功率控制损耗为1 dB,120°扇区天线,天线增益损耗为1 dB。

习题

6.1 (a) 验证模拟AMPS系统的容量(可能容纳的用户数)的数值832,则D-AMPS(IS-136)的相应数值是多少?

(b) 说明欧洲对GSM的频谱分配使992个信道可利用。

(c) 验证表6.2中所表示的AMPS、D-AMPS和GSM系统的容量。

6.2 (a) 用自己的话解释用户信息比特率和传输比特率之间的区别。具体地,说明为什么GSM的用户比特率是22.8 Kbps而传输率是270.833 Kbps。

(b) 计算IS-136(D-AMPS)相应的信息比特率和传输率。

6.3 (a) 画出CDMA系统的示意图并指出其主要的结构模块。

(b) CDMA系统功率控制的主要目的是什么?

(c) 软切换指的是什么?

(d) 解释为什么CDMA可以提供全面的信道重用,即每个信道都可以在所有的小区单元中使用。

6.4 比较TDMA和CDMA方案,指出每一种方案相对另外一种的优势。

---

① Gudmundson, B. *et al*. 1992. "A comparison of CDMA and TDMA systems," Proc. IEEE Vehicular Technology Conference, Denver, CO, May, 732-735.

② Sköld, J. *et al*. 1995. "Performance and characteristics of GSM-based PCS," Proc. 45th IEEE Vehicular Technology Conference, 743-748.

6.5 CDMA 系统 IS-95 语音编码器的比特率是 9.6 Kbps。如果码率是 1.2288 Mcps,则得到扩频增益。这表示每比特有多少码片呢? 在一个比特序列的上面添加码片序列的略图。

6.6 (a) 在文中的假设下,说明美国使用的 CDMA 系统 IS-95,在只考虑一个单独小区干扰的情况下具有容纳 520 个用户的潜力。

(b) 说明美国 IS-95 系统在考虑小区间干扰和文中提到的其他因素的情况下,容量减小到表 6.2 表示的数值范围。

6.7 (a) 使用式(6.5)[或在可用的地方使用式(6.7)]和式(6.8)作图并比较 PSK 误比特率作为信号比特能量和干扰功率谱密度 E_b/I_0 的函数,请分别在衰落和非衰落的环境中比较。

(b) 在数字传输的衰落信道环境中使用任何干扰效应(文中出现了一个例子),检验表 6.1 中的误比特率。

6.8 这个问题用来帮助理解文中用到的一些 CDMA 干扰的表达式。

(a) 说明式(6.13)表示的由于外部移动台面造成的基站的平均干扰功率。

(b) 解释为什么在计算式(6.13)中的阴影衰落表达式时要考虑两个路径。

(c) 根据文中式(6.13)的阴影衰落表达式的计算来说明可以得到式(6.16)。用式(6.16)来计算文中的例子,说明平均小区外部干扰功率在式(6.17)中给出。比较这个结果和阴影衰落标准差 σ 的其他数值。

(d) 用圆形小区干扰模型说明式(6.19b)中第 1 层的干扰功率。

6.9 验证式(6.18)中所表示的包含整个 CDMA 干扰功率的 E_b/I_0 比值,得到对阴影衰落参数 σ 的其他选择值的等同表达式[见习题 6.8(c)]。

6.10 (a) 文中简要讨论了 CDMA 系统的一些像软切换、自动语音静默检测这些能够改善系统容量的特征。为什么这些特征也能够用来补充 TDMA 系统? 解释你的结论。

(b) 把这些特征的效果合在一起,说明每小区 84 个用户的容量是怎么得到的。说明这个数字是怎样随着对不同特征所做的假设的变化而变化的。

第 7 章 检错纠错编码

我们知道，无线传播许多时候都处于比较恶劣的环境，这在第 2 章讨论无线传播环境中的衰落现象时就特别指出了。在第 8 章讨论面向语音的第二代(2G)蜂窝系统时，我们将看到编码技术被广泛地采用，以减小各种传输变异情况的影响，例如，信息传输的衰落等。“编码”意味着有目的地在数字信息流中引入一些附加比特，在接收端通过附加比特来对接收信息流中可能出错的信息比特进行检测或者纠错。在第 10 章介绍第三代(3G)蜂窝系统的分组交换数据传输，以及在第 12 章介绍无线局域网时，都将看到编码技术的广泛应用。在无线系统中，通常采用 3 种编码技术：通过对每个传输的信息加入检测比特进行检错的分组码；纠错的卷积编码；在高速 3G 系统中采用 turbo 码代替卷积编码(turbo 码是一种较新的、高效的编码技术)。虽然分组码在电信系统中广泛应用于纠错，但是在接下来的 7.1 节中将看到，在无线系统中通常采用分组码进行检错，而并没有像卷积码和 turbo 码那样通过纠错程序来进行纠错。在这一章中，我们将对上述的 3 种编码技术进行简明的介绍，希望能让读者更好地理解它们在今后第 8 章和第 10 章所介绍的系统中的应用。本章对分组码和卷积码的讨论与总结，与 Schwartz (1990)[①]在著作中对此进行的详细介绍与论述非常接近。更加全面充分的论述和分析可以参考 Michelson 和 Levesque(1985)[②]的著作。这两本书都包含了这个领域的其他参考文献。本章将从分组码的叙述开始，接着讲述卷积码，最后还将对 turbo 码做简要的介绍。

7.1 检错与纠错的分组码

假设有一块 k 比特的数字信息，分组码编码由 r 个附加比特组成，这 r 个附加比特是通过 k 个信息比特的线性运算得到的。这些附加比特经常称为奇偶校验比特。图 7.1 形象地描绘了这个概念(注意，比特流是由左往右读取，最左边的比特是第 1 个比特)。由 k 和 r 的数值，以及产生奇偶校验位的线性运算，就能够检测或者纠正在通信过程中一个 $n = k + r$ 长的数据块中的错误。这样的码称为一个(n, k)码，而得到的 n 比特编码数据块称为一个码字(块长为 k 比特的信息数据有 2^k 个不同的码字)。

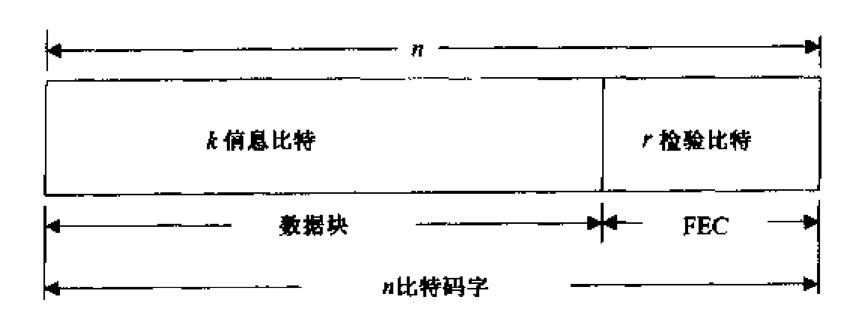


图 7.1 分组码编码

① Schwartz, M. 1990. *Information Transmission, Modulation, and Noise*, 4th edn, New York, McGraw-Hill.

② Michelson, A. M. and A. H. Levesque. 1985. *Error-Control Techniques for Digital Communication*, New York, John Wiley & Sons.

看看下面这一简单的例子。假设 $r=1$ 个奇偶校验信息比特等于所有 k 个信息比特的模 2 和,它附加在这 k 比特数据块的末尾,那么所有 $n=k+1$ 个传输比特的模 2 和总是为 0,而这附加的奇偶校验比特称为一个偶校验比特。很容易想到它能够检测出任意奇数个比特的错误。类似地,若让奇偶校验比特等于所有 k 个信息比特的模 2 和的补,也同样能够得到检测任意奇数个比特错误的奇偶校验方案。例如,让信息块为 00110101,这个 8 比特的数据块从左往右读取(首先是 0,然后是 0,1,1,…)。它的模 2 和为 0,所以偶校验为 0,于是全部 9 比特的码字为 001101010。它的模 2 和的补为 1,所以奇校验为 1。在这两种校验方案下,任意奇数个错误都是能够被检测到的。

现在考虑增加奇偶校验位的数目 r,很显然,这样就能够检测更多的错误并且能够纠正一些错误比特。如何适当地选择奇偶校验比特的产生,例如,如何选择信息比特的线性运算式,是编码理论的基本问题。目前,发展了许多具有不同属性的码,有的码能够检测或者纠正独立出现的零星误码,有的码能够检测或纠正突发误码,有的码能够提供无误或者低误码的同步二进制信息,等等。由于整个序列在传输中被破坏而导致的突发误码,在无线衰落环境下是相当常见的,特别是当缓慢移动的终端处于衰落区域时。然而,如果先把连续的数据块比特存储下来,然后按照某种与其产生的顺序不相同的顺序读取输出,则突发误码的出现就能够减少。这种块交织的技术让原本在同一数据块的连续比特分散开,从而,突发误码就转化成了独立的零星误码(这种技术在蜂窝系统中被普遍应用,在第 8 章和第 10 章中将会提到)。要开发一种具有特定属性的码,重要的结构问题需要重点设计。特别地,在这一编码介绍性的章节中,将主要集中介绍系统码。所谓系统码,即是所有的奇偶校验位都位于码字的末尾,就像前面提到的在单个奇偶校验位的例子中所假设的那样。在本节中,我们还将强调介绍循环码,循环码能够用多项式形式来描述,这个特点使得它的实现和分析都特别容易完成。

考虑现在有一个 k 比特的数据信息块为 $d_1,d_2,\cdots,d_k$,其中 $d_i=0$ 或者 $1,i=1,2,\cdots,k$。用向量 $\boldsymbol{d}$ 来表示这个二进制序列

$$\boldsymbol{d}=(d_1,d_2,\ldots,d_k)$$

它对应的码字用一个 n 比特的向量 $\boldsymbol{c}$ 表示

$$\boldsymbol{c}=(c_1,c_2,\ldots,c_k,c_{k+1},\ldots,c_n)$$

一个系统码则有 $c_1=d_1,c_2=d_2,\cdots,c_k=d_k$,也就是用 $c_1,c_2,\cdots,c_k$ 表示前 k 个传输比特。$r=n-k$ 个奇偶校验比特 $c_{k+1},c_{k+2},\cdots,c_n$ 则由式(7.1)所给出的数据比特的加权模 2 和得到

$$\begin{aligned}
c_{k+1}&=h_{11}d_1\oplus h_{12}d_2\oplus\ldots\oplus h_{1k}d_k\\
c_{k+2}&=h_{21}d_1\oplus h_{22}d_2\oplus\ldots\oplus h_{2k}d_k\\
&\vdots\\
c_n&=h_{r1}d_1\oplus h_{r2}d_2\oplus\ldots\oplus h_{rk}d_k
\end{aligned}\tag{7.1}$$

元素 h_{ij} 取值为 0 或 1,表示上面所说的权重。

举个简单的例子,一个(7,3)码(每输入 $k=3$ 个比特编码器就输出 $n=7$ 个比特),根据下面的运算式得到 4 比特的奇偶校验位:

$$\begin{aligned}
c_4&=d_1\oplus d_3\\
c_5&=d_1\oplus d_2\oplus d_3\\
c_6&=d_1\oplus d_2\\
c_7&=d_2\oplus d_3
\end{aligned}\tag{7.2}$$

于是，结果就有如表 7.1 中所示的 $2^3=8$ 种可能的码字(表中每列的码字，都特意在其输入的信息比特和 4 比特奇偶校验位之间插入了空格)。

表 7.1 码字，(7,3)码

d	c
000	000 0000
001	001 1101
010	010 0111
011	011 1010
100	100 1110
101	101 0011
110	110 1001
111	111 0100

细看式(7.1)以及例子(7,3)码的式(7.2)，都显示出作为一个系统码的码字，由 n 个元素组成的向量 $\boldsymbol{c}$ 可以写成一个 $k\times n$ 的矩阵作用在由 k 个元素组成的信息数据向量 $\boldsymbol{d}$ 上

$$\boldsymbol{c}=\boldsymbol{d}\boldsymbol{G} \tag{7.3}$$

因为 $\boldsymbol{c}$ 是一个系统码字[可以参见例子(7,3)码的码字]，上式中的 $n\times k$ 矩阵 $\boldsymbol{G}$ 称为码生成矩阵，它必须满足前 k 列组成一个 $k\times k$ 的单位矩阵 $\boldsymbol{I}_k$，剩下的 $r=n-k$ 列则是式(7.1)中权系数 h_{ij} 数组的转置(为什么定义 G 为权系数数组的转置？提示：如果选择 $\boldsymbol{c}=\boldsymbol{G}\boldsymbol{d}$ 则结果怎样？)。这 r 列数组由一个 $k\times r$ 矩阵 $\boldsymbol{P}$ 来表示，于是有

$$\boldsymbol{G}=[\boldsymbol{I}_k\boldsymbol{P}] \tag{7.4}$$

其中，

$$\boldsymbol{P}=\begin{bmatrix} h_{11} & h_{21} & \cdots & h_{r1} \\ h_{12} & h_{22} & \cdots & h_{r2} \\ \vdots & \vdots & \vdots & \vdots \\ h_{1k} & h_{2k} & \cdots & h_{rk} \end{bmatrix} \tag{7.5}$$

考虑前面(7,3)码的例子，读者可以自己推导得到

$$\boldsymbol{P}=\begin{bmatrix} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 1 \\ 1 & 1 & 0 & 1 \end{bmatrix} \tag{7.6}$$

以及

$$\boldsymbol{G}=\begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 & 1 & 0 & 1 \end{bmatrix} \tag{7.7}$$

继续看(7,3)码这个例子，仔细看看上表 7.1 中的 8 个码字。注意到每个码字都至少有 3 个比特位置取值不同。一个 8 比特传输码字在任意位置出现单个比特的错误都是可以纠正的，因为接收的码字将会更接近传输的码字，而其他 7 个码字都会与接收码字相比有超过 1 比特的差别。这个(7,3)码就是一个纠正单个差错的码的实例(它还能够检测一定数量的误码，尽管并不能纠正它们。我们将在稍后讨论检错)。这种纠正单个差错码的能力很明显都是归功于它的每个码

字之间至少有 3 比特位置是不相同的这一事实。两个码字之间不相同的比特位置的个数称为汉明距,它是根据一位卓越的研究者 R.W.Hamming 来命名的,因为 R.W.Hamming 在 20 世纪 40 年代到 50 年代之间在贝尔实验室对通信系统的编码理论进行了许多基础性的研究工作。为了能纠正 2 比特误码,汉明距 d 应该为 5;为了能纠正比特误码,汉明距至少应该为 $2x+1$。

给定一个使用的码,怎样对接收的 n 比特码字进行译码呢?一个显而易见的方法就是在接收端重新计算奇偶校验位,并把计算结果与接收的奇偶校验位进行比较。如果它们之间不相同,则表示至少存在一个误码。如果该码是为了纠错而设计的,则误码纠错程序将被调用。一个误码纠错程序的例子就是在码字表中比较所有码字,选出与接收码字最接近的作为传输码字(这个方法对于长码将难以实现,因为码字表的长度随着数据块的长度 k 呈指数增长)。

接收端的奇偶校验过程描述如下:因为 n 比特的码字 $\boldsymbol{c}=\boldsymbol{dG}$ 是一个系统码,它的生成矩阵 $\boldsymbol{G}$ 如式(7.4)所示,很显然码字 $\boldsymbol{c}$ 可以写成

$$\boldsymbol{c}=[\boldsymbol{d}\ \boldsymbol{dP}]=[\boldsymbol{d}\ \boldsymbol{c}_P] \tag{7.8}$$

其中,符号 $\boldsymbol{c}_P$ 代表 $r=n-k$ 比特的奇偶校验位序列。假设接收系统接收到一个 n 比特的码字。它的前 k 比特必须是 2^k 个可能的 k 比特数据向量 $\boldsymbol{d}$ 中的一个。利用这个接收向量 $\boldsymbol{d}$,译码器进行 $\boldsymbol{dP}$ 运算,并且对运算的结果 r 比特向量与接收到的奇偶校验位向量(接收码字的后 $r=n-k$ 比特)逐比特的进行模 2 比较。如果接收序列是一个正确的码字,则比较结果一定为 0。把上述比较数学表达式书写如下:

$$\boldsymbol{dP}\oplus\boldsymbol{c}_P=0 \tag{7.9}$$

(注意,模 2 减等价于模 2 加。)

式(7.9)按照向量–矩阵形式重新书写,可以得到另一等价表示

$$[\boldsymbol{d}\quad \boldsymbol{c}_P]\begin{bmatrix}\boldsymbol{P}\\ \boldsymbol{I}_{n-k}\end{bmatrix}=0 \tag{7.10}$$

符号 I_{n-k} 表示一个阶数为 $n-k$ 的单位矩阵。现在定义一个新的 $n\times r$ 矩阵 $\boldsymbol{H}^T$,由式(7.10)乘以式中后面的那个因子来表示

$$\boldsymbol{H}^T=\begin{bmatrix}\boldsymbol{P}\\ \boldsymbol{I}_{n-k}\end{bmatrix} \tag{7.11}$$

然后,由式(7.8)和式(7.10),得

$$\boldsymbol{cH}^T=0 \tag{7.12}$$

$\boldsymbol{H}^T$ 的转置称为监督矩阵 $\boldsymbol{H}$,于是从式(7.11)中可以得到

$$\boldsymbol{H}=[\boldsymbol{P}^T\quad \boldsymbol{I}_{n-k}] \tag{7.13}$$

监督矩阵在纠错码的研究过程中具有很重要的作用。注意到根据式(7.5)可以得到它的形式为

$$\boldsymbol{H}=\begin{bmatrix} h_{11} & h_{12} & \cdots & h_{1k} & 1 & 0 & \cdots & 0\\ h_{21} & h_{22} & \cdots & h_{2k} & 0 & 1 & \cdots & 0\\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots\\ h_{r1} & h_{r2} & \cdots & h_{rk} & 0 & 0 & \vdots & 1 \end{bmatrix} \tag{7.14}$$

例如,对于前面讨论的(7,3)码,读者可以自己推导得到它的监督矩阵为

$$H=\begin{bmatrix}1&0&1&1&0&0&0\\1&1&1&0&1&0&0\\1&1&0&0&0&1&0\\0&1&1&0&0&0&1\end{bmatrix} \tag{7.15}$$

(在这个例子中,矩阵 H^T 是什么?)

现在回到接收端的检错,以及尽可能的纠错这一问题上。我们指出了式(7.9)提供的一种检测传输过程中误码的方法,而矩阵 H^T 的使用则提供了另一种方法。假设码字向量 c 发生了 1 位或者多位的误码,接收向量 r 则可以写成传输向量 c 与一个 n 比特误差向量 e 的模 2 和,其中误差向量 e 是在发生误码的数位上为 1 而在其他数位为 0 的一个向量。

$$r=c\oplus e \tag{7.16}$$

例如,如果在码字的第 1 位、第 2 位以及最后 1 位发生误码,则有 $e=(1,1,0,0,\cdots,1)$。对接收向量 r 右乘矩阵 H^T,并利用式(7.12)得到

$$rH^T=(c\oplus e)H^T=eH^T\equiv s \tag{7.17}$$

这个由 r 个元素组成的向量 s 称为伴随式。如果它是一个非零向量,则表示有 1 位或者多位误码(反之,零向量 s 并不保证没有误码,因为该码并不一定能完全检测出所有的误码图样)。

作为特殊情况,假设传输码字的第 i 位出现了单个误码,则误差向量 e 只在第 i 位为 1,其余位为 0。由式(7.17)和矩阵 H^T 的定义可知,伴随式 s 一定等于矩阵 H^T 的第 i 行或者说等于监督矩阵 H 的第 i 列。

$$s=(h_{1i},h_{2i},\ldots,h_{ri}) \tag{7.18}$$

当 H^T 的所有行(或者 H 的所有列)都不相同时,误码不仅仅能够被检测,还能够被纠正。满足这个条件的码将能够纠正单个误码。再看(7,3)码这个例子,假设信息向量由 3 比特 110 组成。它的 7 比特的码字根据之前的讨论,应该是 1101001。假设误码发生在第 3 位,由 0 变成 1,则误差向量 e 为(0010000),而接收向量 r 为(1111001)。对 r 右乘矩阵 H^T[见式(7.17)],可以得到伴随式 $s=(1101)$。与式(7.15)比较,注意到它恰好是监督矩阵 H 的第 3 列或者 H^T 的第 3 行。

将上面的讨论一般化,对于一个能够纠正多于 1 个误码的码有什么要求呢?所谓的汉明界(这类界中的一个)提供了计算需要的奇偶校验比特个数的一个方法。假设要至少纠正 t 个误码。汉明界告诉我们,所有可能的校验图样的个数至少等于最多 t 个误码发生的所有可能的情况的个数。对于(n,k)码,$r=n-k$,必须满足

$$2^r\geqslant\sum_{i=0}^{t}\binom{n}{i} \tag{7.19}$$

例如,对于纠正单个差错的码,$t=1$,令 $2^r\geqslant n+1$。如果 $r=3$,则 n 最多不能大于 7 比特。所用的(7,3)码恰好就是这样一个例子,所以它显然属于纠正单个差错的码的范畴。读者自己可以推导(7,4)码和(15,11)码也都能纠正一个错误。但要注意汉明界仅仅如此,它只是给出了纠正某一数量误码的一种可能的码,并不具有确定性。它仅仅提供了一些候选码,却没有提供找到这一确定码的方法。我们将在后面看到,循环码的多项式公式提供了一种构造码的方法。

考虑至少纠正两个错误的可能情况。利用式(7.19),可以找到(10,4)码、(11,5)码、(15,8)码,它们都能纠正两个错误。之前指出的结果码字,汉明距应该为 5。为了纠正至少 3 个错

误,汉明距必须为7。能够纠正至少3个错误码的例子有(15,5)码、(23,12)码、(24,12)码等。这些例子都可以用式(7.19)的汉明界来检验。

然而,在试图用(n,k)分组码进行误码纠错时还存在两个问题。首先,当码字较长时,即使为了纠正较少的误码,也可能必须需要相当多的奇偶校验比特,所以就限制了码字的长度必须在一个较小的范围内。长信息块将导致需要大量的奇偶校验比特。这个问题归结起来就是编码效率的问题。编码效率定义为k/n的比值,它将随着纠错能力的增加而降低。第2个问题是由第1个问题带来的:对于实时传输,传输速率、传输带宽都将会随着奇偶校验比特的增加而增大。传输编码后的$n=k+r$比特必须跟传输编码前的k比特信息块占用相同的时间间隔。这一增加的带宽将会导致系统附加噪声的引入。另一方面,如果令传输原始信息块的时间间隔保持一定,则比特能量将减少为原来的k/n。每比特能量的减少,使得接收码字的误码概率增加了。比特能量E_b与比特差错率的特定关系可参照第6章6.4节中把误码概率与信号能量和噪声谱密度联系起来的几个重要的等式。可以看到,例如对于PSK数字传输,在白噪声以及非衰落环境下,比特差错率为$P_e=\frac{1}{2}\mathrm{erfc}\sqrt{E_b/N_0}$,其中erfc是误差函数的补函数,$E_b$是比特能量,$N_0$是噪声谱密度[见式(6.5)]。由此可见,当误差概率较小时,误差概率随着主要参数E_b/N_0的增大而呈指数减小。然而,在衰落环境下,比特差错率仅仅随着E_b/N_0的增大而线性减小(回顾第6章的讨论,指出在蜂窝环境下,高斯近似的信号干扰取代白噪声,这时差错概率等式依然成立)。所以,当奇偶校验比特增加时,减小比特长度结果使比特能量相应减小,比特差错率增加,抵消了采用该编码原本预想的纠错能力。

为了演示这一效应,考虑在本节作为纠正单个差错码举例的(7,3)码。如果不采用该码,则这3比特未编码数据的差错概率等于3个比特中至少1个发生错误的概率。为简单起见,假设各比特差错是独立的,等于$1-(1-P_e)^3\approx 3P_e$,$P_e\ll 1$。而对于编码的情况,比特差错率等于P_e,但是信号能量则减小为原来的3/7(为什么会这样?)。所以,利用符号p代替这个概率,对于非衰落白噪声(或者"白色干扰")情况,有$p=\frac{1}{2}\mathrm{erfc}\sqrt{3E_b/7N_0}$,因此$p>P_e$,与上面讨论的结果一样(留给读者自己推导在衰落环境下的情况,利用第6章相应的等式)。在编码情况下,呈块误码的可能性等于在7比特中至少有两个独立误码发生的概率(为什么?)。当p较小时,它近似等于$21p^2$。比较这一表达式与未编码情况下的$3P_e$,很明显编码只在p足够小时有效,也就是说,只有当E_b/N_0足够大时,p足够小,才能使2次项p^2起主要作用。细节留给读者自己分析。Schwartz(1990)对(7,4)这一纠正单个差错的码也进行过类似的分析。他的分析指出,对于白噪声非衰落的情况,(7,4)码直至比特差错率P_e小于等于10^{-5}时才使误码纠错得到显著改善。对应的E_b/N_0值为9 dB及以上。所以,在蜂窝环境下,分组码更常用来检错,而本章稍候讨论的卷积码则用来误码纠错。分组码的检错能力可在接下来讨论完循环码后明显地看到。

循环码:多项式表示法

上面讨论的汉明界提供了一种决定码长大小的方法,而码长大小是满足一定误码纠错能力所需的。但是它并没有给出实际构造码的方法,也就是说,没有给出怎样确定奇偶校验比特生成的方法。为了确定某一特定码,则需要额外的结构。循环码是一类能够提供所需结构的码,而且我们将看到,循环码可以用多项式的形式来表示,并允许它们使用移位寄存器技术来

生成。在蜂窝系统中，应用分组码来进行检错而不是纠错，主要是因为它特有的多项式表示式。这将在本节稍候给出一些例子。

循环码定义为具有循环性和线性分组码的封闭性的分组系统码。所谓循环性是指循环码中任意许用码字经过循环移位后得到的码字仍然是一个许用码字；所谓线性分组码的封闭性是指任意两个许用码字的逐比特模 2 和仍然是一个许用码字。所以，如果一个 n 比特码字表示成码向量 $\boldsymbol{c}=(c_1,c_2,\cdots,c_n)$，则码向量 $(c_2c_3,\cdots,c_n,c_1)$ 仍然是一个许用码字。同理，$(c_3,c_4,\cdots,c_n,c_1,c_2)$ 等也是。考虑这一节始终作为例子讨论的(7,3)码。注意，除了全零码字外的所有许用码字都满足上述的循环性。把码字(0011101)和(0100111)逐比特进行模 2 和，结果是(0111010)仍在许用码字中。所以，(7,3)码是一个循环码的例子，并且我们将很快看到它也具有多项式表达式。

循环码的定义使得其对应的码生成矩阵 $\boldsymbol{G}$ 具有一些特殊的性质。于是，用反证法可以得到 $\boldsymbol{G}$ 的最后一个元素，即第 k 行第 n 列的元素一定是 1。假设该元素为 0，让我们考虑信息码字 $\boldsymbol{d}$ 为(000⋯01)的情况。因为相对应的码字 $\boldsymbol{c}$ 由式(7.3)可得 $\boldsymbol{c}=\boldsymbol{dG}$，则必然得到 $\boldsymbol{c}=$(000⋯01⋯0)，即 $\boldsymbol{c}$ 的最后一个比特为 0。将这个码字向右循环移一位，得到一个信息比特为全零的码字，然而它的奇偶校验比特却是 1 的形式。但是，我们知道，根据编码结构这种情况是不可能的。全零的信息序列必然会得到一个全零的奇偶校验序列，所以生成矩阵的最后一行必须以 1 结尾。注意等式(7.7)，一直作为例子讨论的(7,3)码的生成矩阵正好具有这一性质。另外，我们还将看到，生成矩阵的最后一行决定了码结构，所有其他行是其前面行的循环移位或者模 2 和。

现在开始用多项式形式描述循环码。特别地，令一个 n 比特的码字 $\boldsymbol{c}=(c_1,c_2,\cdots,c_n)$，用一个 $(n-1)$ 阶的 x 多项式来表示，其中 x 的每一阶代表 1 比特的移位。于是，最高阶项可写成 c_1x^{n-1}，下一项可写成 c_2x^{n-2}，等等，则最后一项为 c_n。整个多项式表达式如下：

$$c(x)=c_1x^{n-1}+c_2x^{n-2}+\cdots+c_{n-1}x+c_n \tag{7.20}$$

连续移位产生别的码字将通过重复进行 $xc(x)\bmod(x^n+1)$ 运算得到。例如，移位一次，得到

$$c_2x^{n-1}+c_3x^{n-2}+\cdots+c_{n-1}x^2+c_nx+c_1$$

注意到在运算结果中该多项式表达式的系数随之移位，于是提供了另一个码字。再移位一次，得到 $x^2c(x)\bmod(x^n+1)=c_3x^{n-1}+c_4x^{n-2}+\cdots+c_nx^2+c_1x+c_2$。我们将在下面给出怎样由一个给定的循环码 (n,k) 来生成对应的多项式 $c(x)$。但在这之前，首先给出矩阵 $\boldsymbol{G}$，如果以多项式形式表示的话。这个非常简单，首先从第 1 列开始，把值为 1 的元素用 x 适当的幂来代替，而把值为 0 的元素留空。再次考虑(7,3)码的例子，从式(7.7)这个码生成矩阵 $\boldsymbol{G}$ 得到它的多项式表达式为

$$\boldsymbol{G}=\begin{bmatrix} x^6 & - & - & x^3 & x^2 & x & - \\ - & x^5 & - & - & x^2 & x & 1 \\ - & - & x^4 & x^3 & x^2 & - & 1 \end{bmatrix} \tag{7.21}$$

注意式(7.21)的最后一行可以用一个多项式 $g(x)=x^4+x^3+x^2+1$ 来表示。第 1 项的阶数恰好是 $n-k=4$，更一般的情况可以推广得到。从原始矩阵 $\boldsymbol{G}$ 的构造方法可以很清楚地指导它的最后一行一定能够表示成多项式的形式 $x^{n-k}+\cdots+1$（回顾证明循环码生成矩阵 $\boldsymbol{G}$ 的最后一个元素一定是 1）。多项式 $g(x)$ 称为码生成多项式。用符号 r 代替奇偶校验比特的个数 $n-k$，则 $g(x)$ 通过一般化可写成如下形式：

$$g(x)=x^r+g_{r-1}x^{r-1}+g_{r-2}x^{r-2}+\cdots+g_2x^2+g_1x+1 \tag{7.22}$$

其中,系数 g_i,$i=1$ 到 $r-1$,取值为 0 或者 1。正如它的命名所暗示的,一旦生成的多项式确定了,则码以及相应的码字就都确定了。事实上,给定了 $g(x)$,整个生成矩阵 G 就都能够写出来。考虑(7,3)码的例子。首先将 $g(x)$填入最后一行,像式(7.21)那样。然后将 $g(x)$乘以 x,写做 $xg(x)$,即为 $x^5+x^4+x^3+x$,得到矩阵的倒数第 2 行。但是,我们知道对于系统码 G 的前 $k=3$ 列必须是一个单位矩阵 I_k 的形式[见式(7.4)],而移位多项式 $xg(x)$的 x^4 项却破坏了单位矩阵的形式。然而,前面提到循环码的一个性质,两个码字的模 2 和仍然是一个许用码字,所以能够简单地通过对 $xg(x)$加上 $g(x)=x^4+x^3+x^2+1$ 来剔除不需要的 x^4 项,结果得到 x^5+x^2+x+1(读者可以自己检验这一模 2 和的结果)。可见,这就恰好是式(7.21)的第 2 行。它的倒数第 2 行也即第 1 行,都可以再次用同样的方法得到:把上次得到的多项式再次乘以 x,得到新的移位多项式,以检查它是否满足单位矩阵 I_k 的条件,在这里是 I_3。如果不满足,则把最后一行的 $g(x)$加上移位多项式。注意在本例中,$x(x^5+x^2+x+1)=x^6+x^3+x^2+x$,它并没有 x^4 项,所以这一多项式可以直接填到对应的位置上,恰好如式(7.21)所示。于是,通过例子给出了如何利用给定的生成多项式 $g(x)$来得到整个生成矩阵 G。循环码就这样通过生成多项式推导出来。

那么,现在如何能够找出(n,k)循环码的生成多项式 $g(x)$呢? 不加证明地给出下面的结论:生成多项式 $g(x)=x^r+\cdots+1$ 是 x^n+1 的一个因式(Michelson 和 Levesque,1985)。(Michelson 和 Levesque 给出了 $g(x)$是 x^n-1 的一个因式的证明。但是这两个结论在模 2 的代数上是等价的。为了能够与本章整个表述一致,更倾向于使用 +1 的表述)。考虑$(7,k)$这一类码作为例子,其中 $r=7-k$。在这种情况下,需要找出 x^7+1 的因式。留给读者自己推导下面等式

$$x^7+1=(x+1)(x^3+x^2+1)(x^3+x+1) \tag{7.23}$$

于是,用两个具有 x^3 项的多项式中的任一个作为 $g(x)$,就可以得到(7,4)形式的两种可能的码,其中 $r=3$。为了找到之前一直举例的(7,3)码的生成多项式,把第 1 个因式和最后 1 个因式相乘,采用模 2 和,就得到了 $x^4+x^3+x^2+1$,这就是前面得到的生成多项式[见式(7.21)的最后一行]。这里简单地认为因式的乘积仍然是因式。

找出 x^n+1 合适的因式作为 $g(x)$,就能够用前面提到的方法找出对应的生成矩阵 G:把 $g(x)$作为 G 的最后一行。然后,通过对 $g(x)$进行移位以及必要时加上 $g(x)$来得到上一行,再一行一行地重复这一过程。得到了 G 以后,就能够找出所有 2^k 个可能的码字。另一个过程是可以让直接通过生成多项式得到 2^k 个可能码字。注意到,因为根据定义 $g(x)$是阶数为 $r=n-k$ 的多项式,根据模的代数的隐含假设,以及假设 $a(x)$是一个阶数为 $k-1$ 的多项式,就可以把 $n-1$ 阶多项式 $c(x)$写成以下形式:

$$c(x)=a(x)g(x) \tag{7.24}$$

但是,这个有 2^k 个可能的多项式,包括它的退化形式 0 和 1。每一个多项式都对应着一个不同的码字。通过扫描每一个多项式并且把它们模 2 乘以给定的 $g(x)$,就可以生成所有可能的码字多项式。其对应的系数 0 或者 1,则提供了所需要的码字。表 7.2 给出了我们所熟悉的(7,3)码的这一过程。Michelson 和 Levesque(1985)给出了利用 x^3+x+1 作为生成多项式得到(7,4)码的 16 个可能码字的类似例子。注意到该表的所有码字与表 7.1 所示的码字相比,除了顺序上的差别之外,都是完全一致的。

表 7.2 由 $g(x)=x^4+x^3+x^2+1$ 生成的(7,3)码的所有码字

$a(x)$	$a(x)g(x)$	码字
0	0	0000000
1	$x^4+x^3+x^2+1$	0011101
x	$x^5+x^4+x^3+x$	0111010
$x+1$	x^5+x^2+x+1	0100111
x^2	$x^6+x^5+x^4+x^2$	1110100
x^2+1	$x^6+x^5+x^3+1$	1101001
x^2+x	$x^6+x^3+x^2+x$	1001110
x^2+x+1	x^6+x^4+x+1	1010011

在介绍循环码多项式表达式的时候就已经提到,这一多项式表达是使循环码能够很直接地用移位寄存器来实现,下面来演示它。首先把 k 个数据比特 $d_1,d_2,\cdots,d_{k-1},d_k$ 写成多项式的形式

$$d(x)=d_1x^{k-1}+d_2x^{k-2}+\cdots+d_{k-1}x+d_k \tag{7.25}$$

它是一个阶数小于等于 $k-1$ 的多项式(例如,当 $d_1=0,d_2=1$ 时它就变成一个阶数为 $k-2$ 的多项式)。于是,$x^{n-k}d(x)$运算的结果将是一个阶数小于等于 $n-1$ 的多项式。把这一多项式除以阶数为 $r=n-k$ 的生成多项式 $g(x)$,得到一个阶数小于等于 $k-1$ 的多项式 $z(x)$加上一个余式多项式 $r(x)$,如下面式(7.26)所示:

$$\frac{x^{n-k}d(x)}{g(x)}=z(x)+\frac{r(x)}{g(x)} \tag{7.26}$$

现在注意到,在模 2 加法下有 $r(x)+r(x)=0$。所以,如果把被除数写成 $x^{n-k}d(x)+r(x)$,则它一定能够被 $g(x)$整除。从式(7.24)知道,这一结果一定是一个许用码字,所以再次用 r 代替 $n-k$ 有

$$c(x)=a(x)g(x)=x^rd(x)+r(x) \tag{7.27}$$

由于 $x^rd(x)$表示将数据比特左移 r 比特,余项 $r(x)$则一定代表着奇偶校验比特。于是,这些奇偶校验比特就可以通过下面的计算得到

$$r(x)=\text{rem}\frac{x^rd(x)}{g(x)} \tag{7.28}$$

符号 rem 代表余式。再次用 $g(x)=x^4+x^3+x^2+1$ 生成的(7,3)码作为例子,令数据向量为 $\boldsymbol{d}=(111)$,对应的多项式是 $d(x)=x^2+x+1$。然后,读者可以自己计算 $x^rd(x)$除以 $g(x)$,在这里 $r=4$,则余式的结果为 $r(x)=x^2$,它对应的 4 比特奇偶校验位为 0100,于是整个码字就是 1110100。这与表 7.1 和表 7.2 的结果是一致的。

式(7.28)所示的移位寄存编码器如图 7.2 所示。一个正方形盒子表示 1 比特延迟。其中,所有加法都默认为模 2 和的运算。权系数 g_1 到 g_{r-1}是前面式(7.22)定义生成多项式 $g(x)$的系数,它们取值为 0 或 1。k 个输入信息比特在图中标注"输入"的地方输入,此时开关都处于"向上"的位置。注意到这些比特就从编码器的输出端输出,得到 n 比特码字的前 k 位。同时,它们也按顺序进入移位寄存器,按照前面所说的操作进行下去。当 k 个输入信息比特中的最后

一位输出编码器时，开关立刻打到“向下”的位置，然后寄存器中的奇偶校验位比特按照时钟一拍输出一个，直到所有 r 比特都输出完毕。注意到这个过程模 2 运算一直在进行，并且比特 0 序列也在两个开关输入之间用来清除寄存器，在所有奇偶校验位输出完毕后等待下一信息数据的输入。例如，(7,3)码的移位寄存编码器如图 7.3 所示。留给读者自己完成整个编码过程，在 8 种可能的输入数据序列中，选几个来检验一下输出结果的奇偶校验比特是否与表 7.1 和表 7.2 相符。

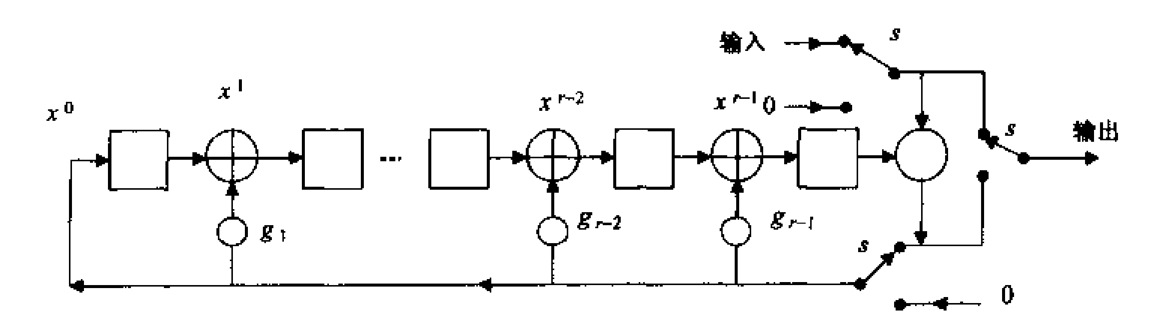


图 7.2 循环码移位寄存编码器，r 比特奇偶校验位。开关向上连接信息比特，向下连接最后 r 比特奇偶校验位

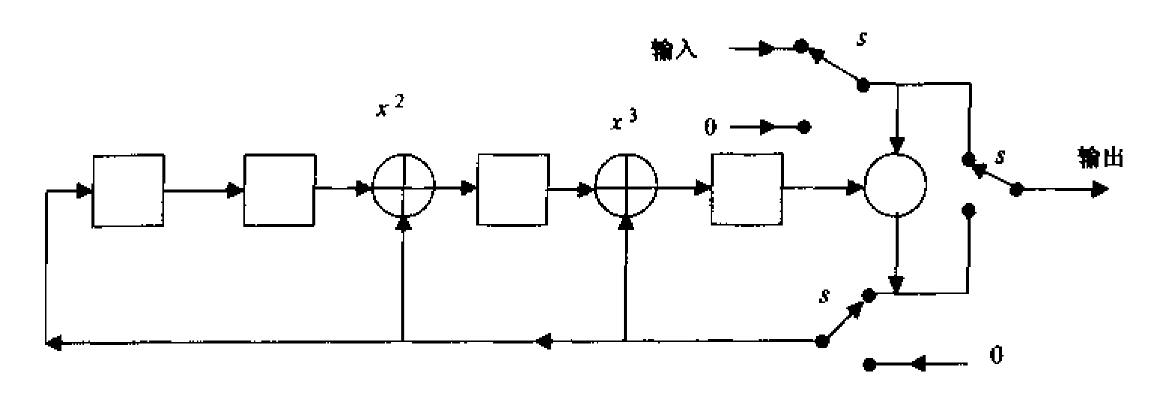


图 7.3 (7,3)码编码器，$g(x)=x^4+x^3+x^2+1$

循环码的检错

根据上面的讨论，我们很清楚地看到循环码的多项式表达式，使得它能够直接用移位寄存器来实现编码。接收系统也进行同样的移位寄存操作，然后将得到的奇偶校验比特与实际接收到的奇偶校验比特进行比较，从而确定在传输过程中是否发生误码。上面所说的重点在于怎样使用这些码来纠正误码。然而，无线蜂窝系统使用分组码的首要目的是检错而不是纠错，这一点在本章前面已经提到。在无线数字系统中，采用的是相当大规模的信息块，在第 8 章和第 10 章的无线系统的举例中将看到它的信息块长度将达到数百比特，所以它排除了使用本章一直讨论的基于分组码的纠错技术。即使是非常有限的误码，如此大的数据块也需要大量的奇偶校验比特来纠正，故奇偶校验只用来检错。像因特网等分组交换的网络也是用奇偶校验计算来进行检错的。在无线系统中，经常使用 8，12 或 16 个奇偶校验比特附加到信息数据块上进行检错。在第 8 章和第 10 章的系统中，我们将看到发送端在进行了奇偶校验分组编码后再进行卷积编码和 turbo 编码来提供误码纠错的能力(在接收端则执行相反的过程：首先卷积码和 turbo 码译码器用来进行误码纠错，然后奇偶校验的分组码则用来检测是否存在尚未被纠

正的误码)。这些码将在下面的章节中讨论。所以,比特奇偶校验检错是用来检测尚未被卷积码纠正的差错。从这一点来看,集中讨论码的多项式表达式来进行检错非常有用。分组码因此而称为外码;而卷积码和 turbo 码则被称为内码。

在讨论循环码检错的应用之前,还有一点非常重要,即为什么在通信系统中是检错而不是之前一直讨论的前向纠错起着非常重要的作用。我们知道在电信系统中有几种可能需要进行检错的情况:当信息数据块中出现 1 个或者多个不能被纠正的误码时,它们被简单地丢弃;接收系统要发送信号通知已经检测出 1 个或者多个误码,并要求进一步采取行动;或者一个可能的行动是接收系统通知发送方重新传送该数据块。第 1 种情况最简单,可以应用在实时业务中,例如语音,在这种情况下接收系统检测到语音数据包发生了 1 个误码就简单地丢弃该包。众所周知,在实时语音中高达 5% 的差错率都是可以容忍的,也就是说,进行语音通话的人在偶尔丢失语音数据包的情况下并不会感觉到通话质量受到明显的影响。下一章节即第 8 章要讨论的第二代数字无线系统就是专门为处理语音通信而设计的。

第 10 章将要讨论的第三代移动通信系统则是为处理分组数据业务而设计的。大多数这种类型的通信不太要求实时传输。在这种情况下,分组重传就可以用来纠正检测到的任意错误。特别地,有一种称为 ARQ(automatic repeat request)的技术,即自动请求重传,被普遍地应用在分组交换网络中,包括因特网,它也被分组交换的无线系统所采用。在这一技术中,发送端在发送一个数据包时设置一个计时器。如果接收端检测无误,它就返回一个肯定的确认,发送方要在计时器超时前收到这一确认。另一方面,如果检测到一个误码,接收方就简单地丢弃该包。发送端的计时器超时,则它就立刻重新发送该数据包,而计时器也重新复位。在因特网的情况下,传输层是在两个端用户之间建立端到端的通信,执行 ARQ 的过程。接收端用户检测所收到数据包的误码情况,并且在无误时向发送端用户回复确认,否则它就丢弃被认为是误码的数据包。当发送端用户在预定的时间内仍没有收到确认时,则重新发送该数据包。这就是为什么许多"底层"的协议,例如在局域网中广泛采用的以太网,只需要检错,而误码纠错则留给了"高层"的协议如 TCP 来完成(将在第 10 章简要地介绍分组交换数字通信中所采用的分层协议的概念)。

一个数据包的检错过程与误码纠错的过程很相似,就像前面提到的一样:计算一系列奇偶校验比特,大多数情况下都使用生成多项式移位寄存器来计算,然后添加到数据包末尾进行发送。接收端进行同样的计算,看看计算结果与接收到的奇偶校验比特是否一致。附加到数据包末尾的奇偶校验比特有很多种不同的叫法,例如分组校验(block check)、校验和比特(checksum bits)、检错域(error-detection field)、帧检测序列(FCS,frame check sequence)和循环冗余校验(CRC,cyclic redundancy checking),以及其他各种名称。在这一章的开始,我们就已经看到,单比特的奇偶校验位能够检测接收信息中所有奇数比特的错误。然而,更一般地,对于广泛地使用奇偶校验编码来进行检错的关键,强调以下结论:采用 r 比特奇偶校验位能够检测误码长度小于等于 r 的所有突发误码,这与信息数据块的长度无关。这一结论的建设性证明可参阅 Schwartz(1990)的著作。注意,一个长度为 b 的突发误码是指接收的误码从第 1 个错误到最后 1 个错误之间有 b 比特,而它们之间的比特可以是正确的或者是错误的。所以,用 8 比特奇偶校验位就能够检测任意长度小于等于 8 比特的突发误码;用 16 比特奇偶校验位就能够检测任意长度小于等于 16 比特的突发误码。

更有趣的是,循环码用 r 比特奇偶校验位能够检测到误码长度大于 r 的突发误码中的

很大一部分。特别地，前面也已经说明，当使用多项式表达的循环码，突发误码长度 b 大于 r 比特时，不能被检测到的比例只占 2^{-r}（Schwartz, 1990）。这一结果与信息块的长度无关，所以，长的数据块可以被使用，并且可以检测更多的错误图样。例如，如果附加了 $r=8$ 个奇偶校验比特，则不能被检测的突发误码图样只占全部的 $2^{-8}=1/256$；如果用 12 个奇偶校验比特，则不能被检测的突发误码图样所占比例则下降到 $2^{-12}=0.000\ 24$；如果用 16 个奇偶校验比特，则不能被检测的突发误码图样所占比例则下降到 0.000 015！注意，在第 8 章和第 10 章我们将看到，上面这些奇偶校验比特的个数，在数字无线系统的空中接口中进行传输信息块时被普遍地使用。

例如，16 比特、12 比特和 8 比特奇偶校验序列就被选用在第 10 章介绍的第三代 cdma2000 无线蜂窝系统的数字与信号信息传输中。系统采用帧质量指示器（Frame Quality Indicator）来分配奇偶校验序列的长度。这 3 种情况的生成多项式由下面式（7.29）到式（7.31）来表示。对应的移位寄存器的实现框图，都是基于图 7.2 的一般实现框图，分别如图 7.4 到图 7.6 所示（cdma, 2002）[①]。

$$g(x)=x^{16}+x^{15}+x^{14}+x^{11}+x^{6}+x^{5}+x^{2}+x+1 \tag{7.29}$$

$$g(x)=x^{12}+x^{11}+x^{10}+x^{9}+x^{8}+x^{4}+x+1 \tag{7.30}$$

$$g(x)=x^{8}+x^{7}+x^{4}+x^{3}+x+1 \tag{7.31}$$

利用前面提到的未检测比例的突发误码比例为 2^{-r} 这一事实，我们可以用"back-of-the envelope"方法来计算 n 比特长的信息块中未被检测的误码的概率。特别地，假设误码模型是块误码概率，并与块长 n 成正比。于是，有块误码概率为 np，其中 p 是一个很小的参数[参见 Schwartz(1990)对该模型的讨论。通过观察可以得到，如果比特差错是独立的，并以 p 为概率发生，当 $np \ll 1$ 时，一个长为 n 的数据块发生错误的概率为 $1-(1-p)^{n}\approx np$。在存在衰落的蜂窝环境下，比特差错通常不是独立的。事实上，衰落突发误码是很常见的。然而，在第 8 章和第 10 章讨论第二代和第三代系统时我们将看到，普遍采用块交织使得每一个数据块的比特在时间上被分散，于是就把可能的突发误码转化成准独立的比特误码。尽管存在这一缺陷，对于块误码概率与块长度成正比这一假设从直观上来看仍然是合理的，它是从 Schwartz(1990) 提到的早期电话信道的一些实验中提出的]。利用这一模型，对于一个长为 n 比特的数据块，其未被检测的误码概率 P_{bl} 具有以下简单的表达式：

$$P_{bl} \approx np2^{-r} \tag{7.32}$$

思考一些例子：令 $p=10^{-5}$。如果 $n=500$ 比特，当采用 $r=8$ 比特奇偶校验位时有 $P_{bl}\approx 2\times10^{-5}$，而采用 16 比特奇偶校验位时有 $P_{bl}\approx10^{-7}$。如果 n 增加到 1000 比特，以上结果显然变成原来的两倍，然而，误码概率仍然非常小。注意，上述结果均是由于选用了小的参数 p 才得到的。

这些奇偶校验检错的讨论结束了对分组码特别是对循环码及其多项式表达式的讨论。从下一节开始，将转向无线系统中普遍使用的卷积码的讨论。

---

① cdma 2002. Physical layer standard for cdma2000 spread spectrum systems, Release A, 3GPP2 C.S2002-A, Version 6.0, 3rd Generation Partnership Project 2, 3GPP2, February, http://www.3gpp2.org/public_html/specs/

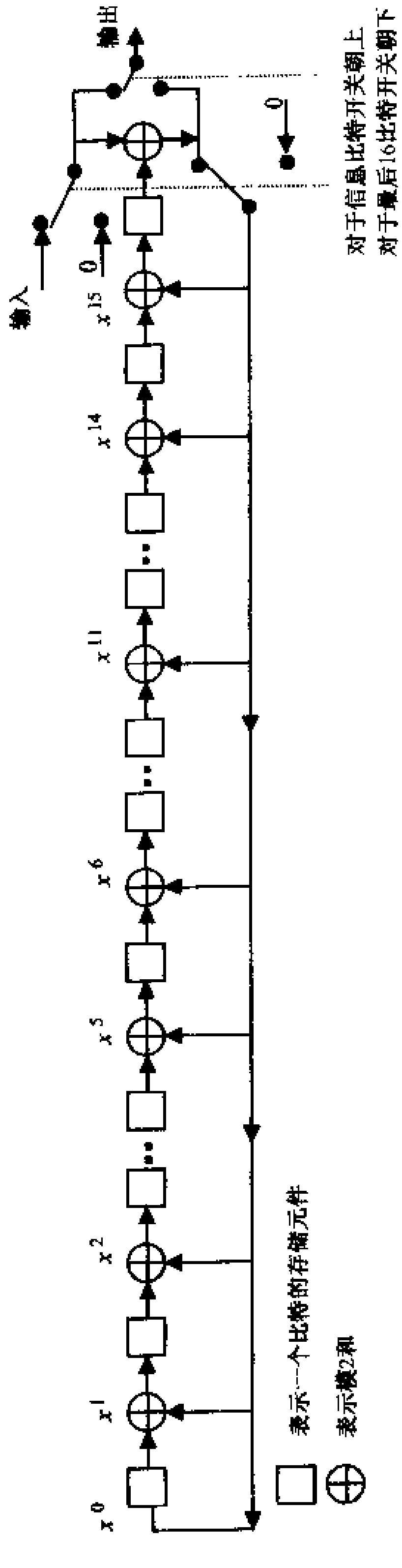


图7.4 帧质量指示为16比特，cdma2000（引自cdma2000：Fig. 2.1.3.1.4.2.1-1）

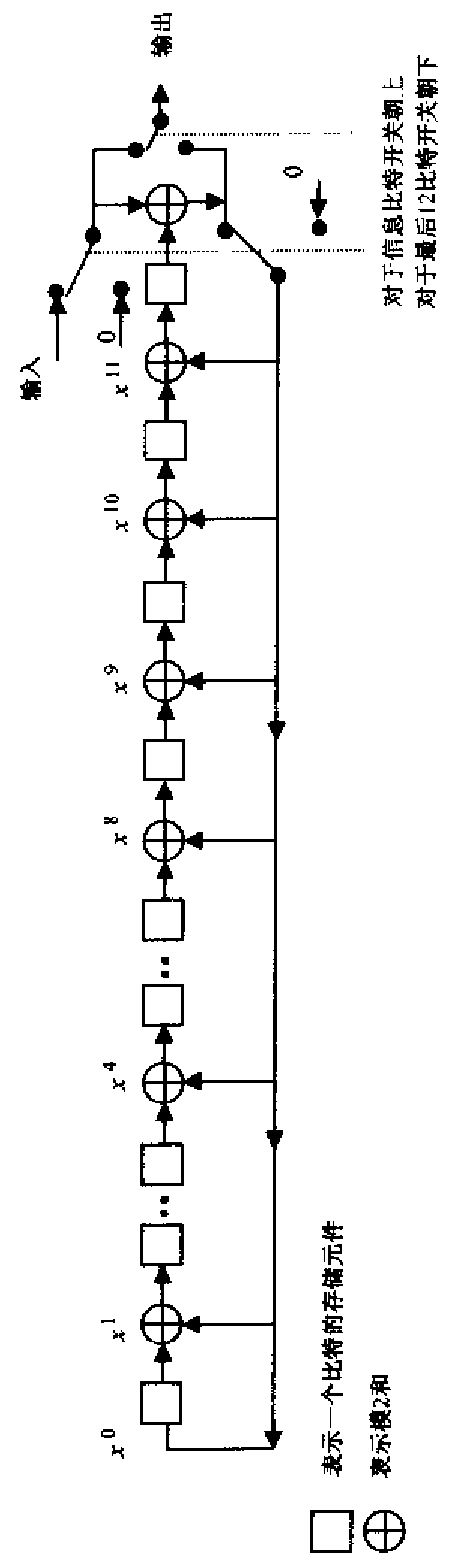


图7.5　帧质量指示为12比特，cdma2000（引自cdma2000：Fig.2.1.3.1.4.2.1-2）

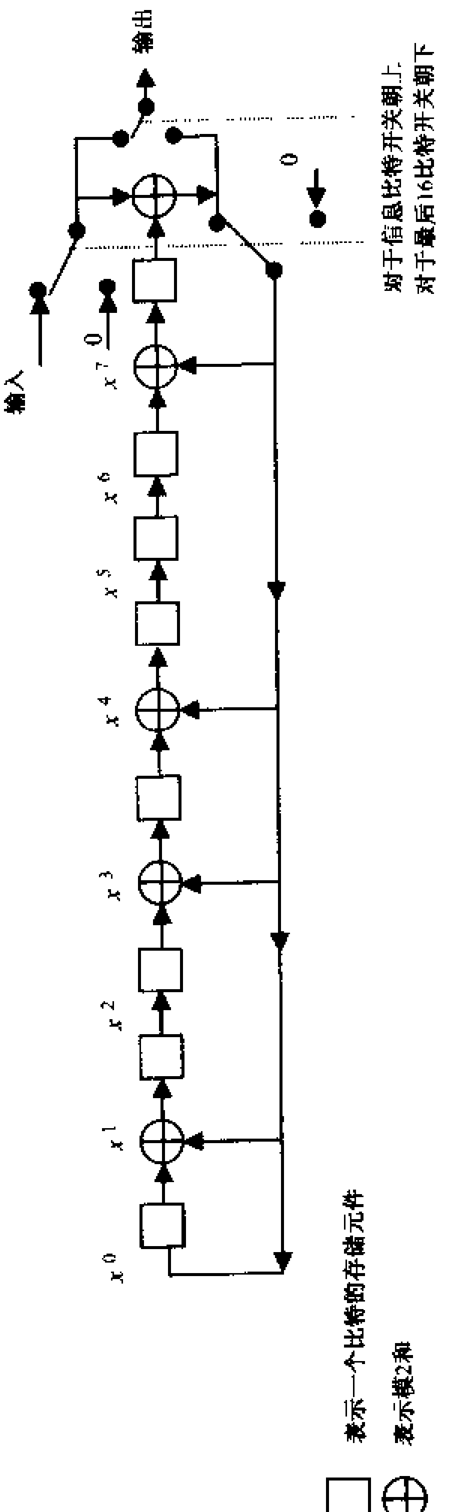


图7.6　帧质量指示为8比特，cdma2000（引自cdma2000：Fig.2.1.3.1.4.2.1-3）

7.2　卷积编码

在前面我们就注意到编码技术可以分为两个主要的类别：分组码和卷积码(下一节所简要讨论的Turbo码是建立在卷积码基础上的)。分组码是在 k 比特信息位后添加 r 比特奇偶校验位，构成一个 n 比特的码字。而对于卷积码，则是在一个滑动的数据比特序列上进行模2和操作，从而生成一个比特码流。卷积码具有误码纠错的能力，它首先被引入卫星和太空的通信中。在接下来的章节中我们还可以看到数字蜂窝系统现在也开始使用这种码。本节所讨论的卷积码很大程度上均引用自 Schwartz(1990)的著作。更多深入的推导与处理可参阅 Michelson 和 Levesque(1985)以及 Proakis(1995)[①]的著作。在介绍的过程中也会引用一些这个领域的基本论文。

循环码由 K 阶移位寄存器组成。输入数据比特每次沿着寄存器移位1比特。将移位寄存器的内容按照一定的关系进行模2和后，以输入速度的 v 倍往外读出，从而达到每输入1比特输出 v 比特的目的。K 称为卷积码的约束长度，这样一个卷积码由约束长度 K 和编码效率 $1/v$ 来分类。稍后将看到，在讨论卷积码的纠错性能时，性能的提高可以通过增加 K 和码率 v 来实现。然而，共同带来的代价是复杂度和运算的比特速率的增加，例如，一个 $K=3$，码率为1/2的编码器，如图7.7所示。所以，输出比特速率是输入的两倍。图中的方格代表寄存器的阶，与图7.2到图7.6类似，图中展示了两种可能的编码器示例。其中，图(a)有 $K=3$ 个方格；而图(b)只用了 $K-1=2$ 个方格。需要注意，在这两种表示中，输入比特都供向输出线，以输入线路2倍的速率输出。它们与其他寄存器的比特按照指定的关系进行模2和。编码输出首先来自输出线路1，然后来自输出线路2，接着再回到1，循环反复，这一过程通过开关在两个输出端之间来回闭合来实现。当比特进入寄存器后，它们就沿着寄存器一次移位1比特。图7.7(b)中只有 $K-1$ 个延时阶在编码器中出现，这是因为它的输出比特是由 $K-1=2$ 阶延时寄存器的输出加上新输入的比特组合而成的。卷积码编码器必须在开始处理输入字之前把所有的寄存器清零，这使得编码结果给每个输入字增加了 K 比特，在第8章和第10章特定的例子中将会清楚地看到这一点。

现在考虑将图7.7的编码器操作作为例子。假设它已经被初始化，并且首先输入的信息比特是一个0，然后很明显输出会有两个0。现在让一个1输入，从图7.7的连接关系很显然得到两个输出是1和1，让第2个1进入编码器。留给读者自己推算此时的两个输出，按照顺序为0 1。现在输入一个0，对应的输出还是0 1。比特流继续沿着 $K-1$ 阶移位寄存器进行操作。对于一个特定的8比特输入序列得到的全部16比特的输出如下：

输入比特： 0　　1　　1　　0　　1　　0　　0　　1
输出比特： 0 0 1 1 0 1 0 1 0 0 1 0 1 1 1 1

注意到不同延时阶的比特按照特定关系进行模2和得到输出，是可以用两个函数生成式 g_1 和 g_2 来表示的，如图7.7所标注的那样。每个函数生成式由一个 K 比特向量来表示，一个比特对应一个移位寄存器，包括输入比特。某一阶对应1，则说明它是加入模2和的运算的；而对应0，则说明不加入运算。看图7.7的例子，有 $g_1=[1\ 1\ 1]$，$g_2=[1\ 0\ 1]$。这些函数生成式用来代替卷积编码器，就像分组码如式(7.22)和图7.2所示的生成多项式一样，用来表示一

① Proakis, J. G. 1995. *Digital Communications*, 3rd edn, New York, McGraw-Hill.

个特定的分组码的寄存器形式。一般地，一个约束长度为 K，码率为 $1/v$ 的编码器有 v 个这样的函数生成式。每一个对应一个输出线，每一个 K 比特向量则指出一个输出比特是由各个移位寄存器按照某种关系进行模 2 和得到的。

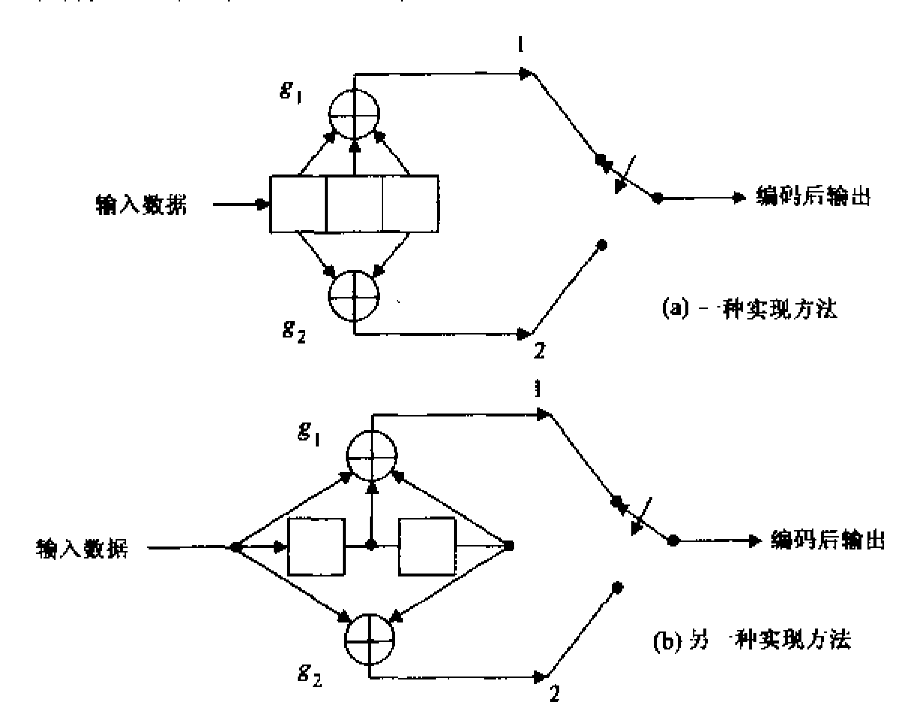


图 7.7　$K=3$，码率为 1/2 的卷积码编码器举例

更一般地，特别是对于约束长度较长的卷积编码器，函数生成式简单地用十进制形式来表示。所以，在图 7.7 的例子中，两个生成式分别可以表示为 $g_1=7, g_2=5$。对于第 10 章将要介绍的第三代 cdma2000 系统，其 $K=9$，码率为 1/2 的卷积码编码器的例子如图 7.8 所示(cdma, 2000)。两列编码输出比特，c_1 和 c_2 被定义为编码输出符号，如图所示。两个对应的函数生成式的十进制表示为 $g_1=753, g_2=561$。可以检验，由图 7.8 中 $K=9$，码率为 1/2 的卷积码编码器的比特连接关系可以得到比特连接向量分别为[111101011]和[101110001]。将它们每 3 个为一组可得到十进制数字，即为 753 和 561。改变连接关系，也就是说改变函数生成式，输出比特流也随之改变。特定的函数生成式的选择决定了卷积编码器的差错概率的性能表现。我们将简要地总结卷积编码器的性能特点。

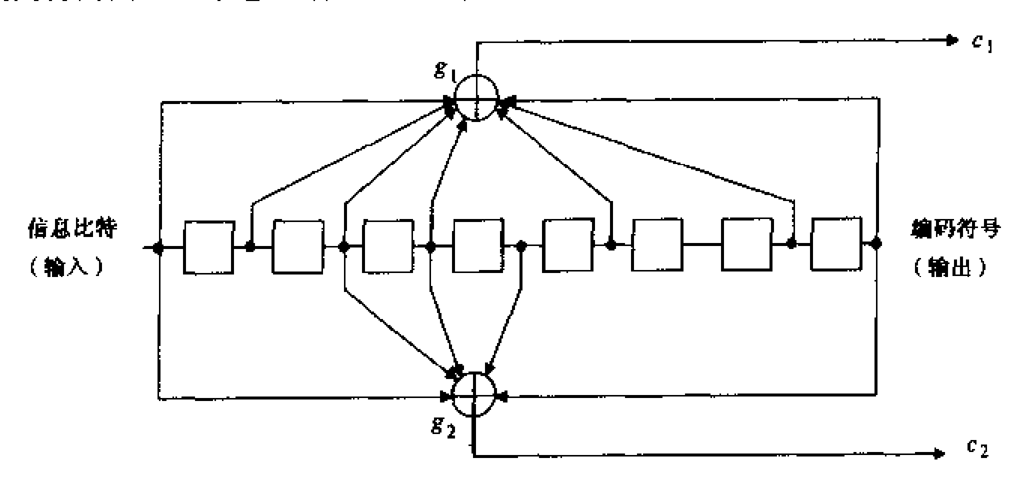


图 7.8　$K=9$，码率为 1/2 的卷积码编码器

一个特定的卷积码可以有 3 种不同的表示形式，状态表、网格图和树状图。这些表示给编码器的操作提供了更好的理解，并且能够直接导出它们的性能特点。以图 7.7 的 $K=3$，码率为 1/2 的卷积码为例来分别展示这 3 种表示方式。首先从状态表开始。回到图 7.7，注意到每个输入比特时间间隔内的输出比特序列是由该时刻输入比特与 $K-1=2$ 个最右边的寄存器[见图 7.7(a)]或者说 $K-1=2$ 个寄存器[见图 7.7(b)]之和来决定的。于是，这 $K-1=2$ 个寄存器的比特有 $2^{K-1}=4$ 种可能的组合，这些比特的可能组合称为状态。例如，图 7.7 的编码器输入一个 0，此时两个寄存器存储的比特为 00，就像前面给出的 8 比特输入/16 比特输出的例子一样。可以说，编码器在 0 到达的时候状态为 00，输出比特也是 00。然而，让一个 1 到达输入端，则得到的输出序列为 11，因为选定的关系是模 2 和。若一个 1 到达输入端，此时 $K-1=2$ 个寄存器的值为 10，可以说，编码器转移到新的状态 10。若一个 0 到达输入端，则编码器的状态停留在 00。图 7.9 给出了图 7.7 的卷积码编码器的状态图。这 $2^{K-1}=4$ 个状态由图中长方形格子中的 $K-1=2$ 比特来表示，并且用 a,b,c,d 来标记，如图所示。带箭头的线表示从一个状态到另一个状态的转移，由括号中显示的输入比特值来决定，并在线的旁边、输入比特之后标注输出比特。所以，如图所示，当编码器处于状态 00 或者说状态 a 时，输入 0 则编码器保持在同一状态，对应输出比特为 00。然而，如果输入 1，则编码器状态转移到 10 或者说状态 b，对应输出序列为 11。当编码器处于状态 b 时，输入 1 则使编码器转移到状态 11 或者说状态 d，对应输出为 01。而如果相反，编码器在状态 b 时输入的是 0，则编码器状态转移到 01 或者说状态 c，对应输出序列为 10。图 7.9 中剩下的状态转移将留给读者自己推导演示。显然，具有不同的函数生成式或者说码连接关系，对应的输出比特序列则不同，但是图 7.9 的状态转移图是一样的。所以，它被应用在所 $K=3$，码率为 1/2 的卷积码编码器中。一个 $K=4$，码率为 1/2 的卷积码编码器有 8 个状态。注意这个状态转移图是循环转移的，并且所有状态都访问过：任意时刻从任意状态出发，卷积码的有限状态机表示法将最终回到该状态。

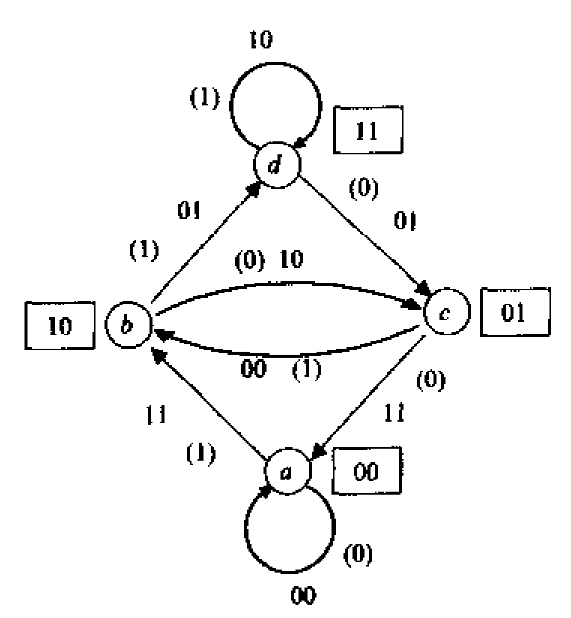


图 7.9　图 7.7 的卷积码编码器的状态转移图

现在开始讨论同一个 $K=3$，码率为 1/2 的卷积码编码器的网格图，如图 7.10 所示。其他任意的卷积码编码器的网格图表示都可以用同样的方法简单地得到。在这个例子中有

$2^{k-1}=4$个状态，它们被排列在垂直的轴上，如图所示。时间，即输入比特的时间间隔，用水平轴表示。状态转移则表示在从一个时间间隔到下一个时间间隔上，由于输入比特为 0 而导致的状态转移通常标注在两个输出状态线中上面的那条线。输出比特按顺序标注在每个状态转移线上，就像图 7.9 中的状态转移图那样。注意网格图的表示法是如何立刻表现出编码器有限状态机循环反复的属性的：从任意状态开始，可以大致观察，看状态之间是如何转换，然后最终回到所关注的状态的。这种表示方法在评价一个给定的卷积编码器的性能时非常有用。

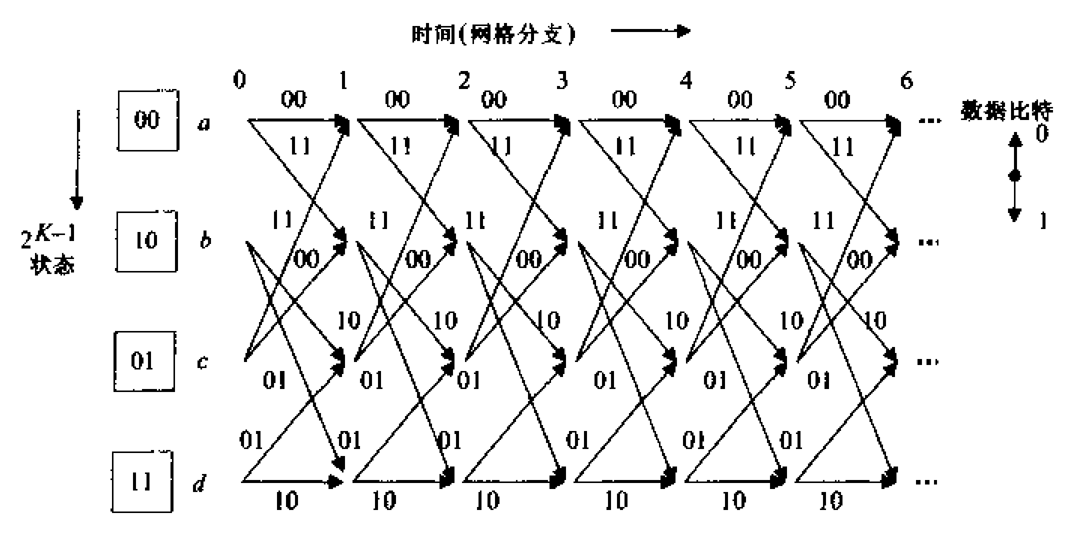


图 7.10 图 7.7 的卷积码编码器的网格图

卷积码编码器的第 3 种表示方法是树状图。再次用同一个 $K=3$，码率为 1/2 的卷积码编码器作为例子对它进行描述。这个例子的树状图表示见图 7.11。它把状态转变的序列独立分开并认为是不同的路径，这种表示方法与网格图一样有效。就像刚才描述图 7.10 的网格图一样，在图 7.11 中的两条分支中，上面的那个对应数据 0 的输入，下面的那个对应数据 1 的输入。对应某一个给定路径的每一个特定的状态转移输出比特，则标注在该状态转移线的上方。这种表示方法过于充分地展示了卷积码编码器状态机的重复性，而显然当跟踪每一个不同的路径时，随着时间的推移，一直就只有 $2^{K-1}=4$ 个状态在进入和再进入。需要注意的重要一点是，每一个可能的路径对应一个不同的输出比特序列，它的数量是随着时间呈指数增长的。而任意一对路径的输出比特具有不相同的数量，即汉明距也是随着时间增长的。这一现象指明了为什么卷积码能够提供较好的性能。当一个指定路径长度不断增长时，它就能更加清楚地与其他路径区分开。在传输过程中，出现的一个或者多个错误就更加有可能被检测或者纠正。这一简单的现象同样也指明了为什么当编码器的约束长度 K 或者码率 v 增加时能够改善系统的性能。

给定一个特定的二进制数据序列进入卷积码编码器，终端的接收系统是如何恢复或者译码从而得到该序列的呢？这一章的主题是带有信息的二进制序列的编码，这意味着某些传输信息在传输过程中因为干扰、衰落或者噪声等原因造成了破坏。所讨论的卷积编码的整个目的就是尽可能的恢复被破坏的信息，以达到最小的差错概率。这个目的定义了“最好”的译码器的概念。有了卷积编码这一方法，给定一个接收到的信息序列，就找到了与接收到的序列对应的最大似然序列。例如，假设选择一个长为 L 个时隙的二进制输入序列。接收序列就有 2^L 个可能的路径。其中，与该序列最相似的一个序列被选择。一个实现这一目的的译码器称为

最大似然译码器(Maximum-likelihood decoder)(Schwartz,1990;Proakis,1995)。考虑如下这样一个译码器的简单版本。对于每一个输入的二进制数据时隙,就有 v 比特信号被传输和 v 比特信号被接收。因为衰落、噪声或者干扰,每个接收信号本质上是模拟形式的,即使传输的是一个二进制的 1 或者 0。为简单起见,假设这 v 比特信号被独立地转译成最好的估计 1 或者 0。这样的技术称为硬限制的译码或者硬判决的译码。这一过程简化了所考虑的讨论[接收信号是真实的模拟信号的最大似然译码在 Schwartz(1990)和 Proakis(1995)的著作中讨论过。研究表明,采用硬判决译码比保留其模拟信号并直接运用它要多付出 2 dB 性能损失的代价(Heller 和 Jacobs,1971)①。这一性能损失是由于要达到相同的差错概率,从而增加信号功率而造成的。重新对接收信号用多于 1 比特的信号来表示,并且用这一表示形式来进行译码则被称为软判决译码。下面所描述的实现译码过程的维特比算法(Viterbi algorithm)就是一种软判决译码(Michelson 和 Levesque,1985;Proakis,1995),所以它能够提供近似最优的性能。让我们继续讨论硬判决译码来简化分析过程]。现在接收路径由一组 L 个符号组成,每个符号包含 v 个比特。接收的 Lv 比特序列现在与 2^L 条可能的路径中的每一个进行比较,汉明距最接近的那一条路经被选择为最似然或者最有可能被传输的路径。对应这条路径的 L 比特则被认为是最可能被传输的数据比特。例如,考虑图 7.7 中 $K=3$,码率为 1/2 的卷积码编码器的树状图如图 7.11 所示。对于 $L=3$,在第 3 个时隙有 8 条可能路径与接收序列进行比较;$L=4$,在第 4 个时隙有 16 条可能路径与接收序列进行比较。走到第 5 个时隙,如果此时要决定在这 5 个时隙里面汉明距最接近接收序列的那条路径,则有 32 条路径需要与接收序列进行比较。

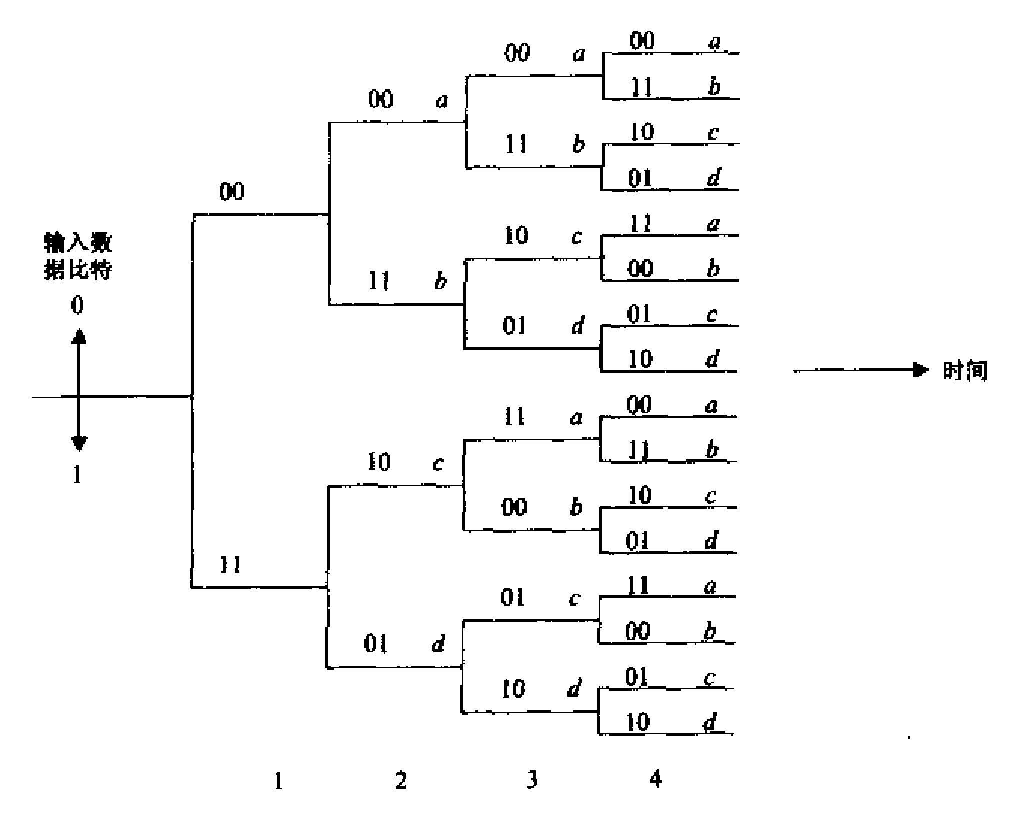


图 7.11 图 7.7 的卷积码编码器的树状图

① Heller, J. A. and I. M. Jacobs. 1971. "Viterbi decoding for space and satellite communications," *IEEE Transactions on Communication Technology*, COM-19, 5, part II (October), 835 – 848.

增大 L 可以降低差错概率，也就是说降低了选择错误路径的概率。但是，注意到增大 L 将带来一个主要问题。当增大 L 时，需要比较的可能路径的数目呈指数增加到 2^L。例如，如果走到 $L=100$，也就是说需要比较长为 100 个时隙的路径的汉明距，则需要 2^{100} 次比较！这个将如何完成？一个具有创造性的最大似然译码算法由 Andrew J. Viterbi 发明，也常称为维特比算法(Viterbi Algorithm)，它提供了这个问题的重要缓和方案，并且已经在卫星和太空通信以及无线应用中被广泛使用(Viterbi, 1967)①。维特比算法的使用前提很简单：要求已经讨论过的卷积码编码器输入的二进制数据有 2^{K-1} 个状态。假设一个给定的二进制时隙处于这些状态中的一个状态，如果要达到这个状态，它的上一个时隙的状态只能有两个可能的状态可以选择。$K=3$，码率为 1/2 的卷积码就是这样一个例子。看图 7.9 的状态表，假设现在在状态 b，则它的前一时刻只能由两种可能状态 c 或者 a。输入二进制为 1，对应编码器输出为 00，将使得状态从 c 转移到现在的状态 b；输入二进制为 1，对应编码器输出为 11，将使得编码器从状态 a 转移到现在的状态 b。假设实际接收的二进制序列与从状态 c 转移到状态 b 的路径的汉明距小于与从状态 a 转移到状态 b 的路径的汉明距，则在下一个时刻，来自状态 a 的路径的汉明距下降到小于来自状态 c 的路径的汉明距是不可能出现的。所以，我们就没有理由保留来自状态 a 的路径，可以丢弃它而保留来自状态 c 的路径，也可以再对每一个编码器的状态进行同样的比较，这就是维特比算法的本质。每经过一个二进制时隙，从 2^{K-1} 个状态中的每一个状态选出与接收序列相比两条路径中汉明距较小的一条，然后保留下来，称为"幸存路径"。所以，在每个二进制时隙，只需要进行 2^{K-1} 次比较，经过 L 个时隙，维特比算法总共进行了 $2^{K-1}L$ 次比较，而不是像标准的最大似然过程那样需要 2^L 次比较。

更一般地，如果不是采用硬判决译码，而是将接收的模拟信息在每个二进制时隙采样下来进行译码，则我们需要用欧几里得距(几何距)来代替汉明距，用它来度量决定每个时隙的幸存者(Schwartz, 1990; Proakis, 1995)。在这种情况下，度量最小的幸存下来，与硬判决译码时采用汉明距一样。路径度量每个时隙都在不断增大。正如上面的附加说明所说，维特比译码器根据一定的开销采用软判决译码，在实际中得到了广泛的接受。

L 应该选用多大的值呢？很显然，总是希望能够让 L 越小越好。研究表明，L 取值为 K 的 4 倍或 5 倍就足以得到与选用相当大的 L 值很接近的性能，也就是说，这是最好的可能选择(Heller 和 Jacobs, 1971)。同样的研究也表明了一个简单的译码流程是对最似然的幸存路径输出最早的输入比特。每一个时隙之后，即在选择了 2^{K-1} 个幸存者后，所有幸存者中度量最小的那条路径，其最早的输入比特被选为与发送端编码器的输出最似然的结果；而所有幸存路径中最早的一个时隙则从路径寄存器中删除掉。在下一时隙重复这一算法。考虑一个像图 7.8 所示 $K=9$ 的编码器作为例子。如果选择 $L=4K$，就有 $L=36$。每一个输出比特需要的比较次数为 $2^{K-1}L=256\times 36=9216$ 而不是 2^{36} 次，这是一个很大的约简！

维特比算法操作过程的一个很简单的例子在表 7.3 中展示，为了简单，还是假设硬判决译码流程。这个例子还是采用图 7.7 所示的 $K=3$，码率为 1/2 的卷积码编码器作为例子。为了简单，假定每隔 5 个输入比特时隙就产生一个判决，且系统已经运行了一些时间。上一组 5 个时隙接收的比特序列(每一个输入比特时隙两个比特)如表 7.3 所示，并且与 8 个可能路径进

① Viterbi, A. J. 1967. "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, IT-13(April), 260-269.

行了比较:来自前一时隙的4个幸存者,由于在进入现在时隙的4个状态中各有两种可能的状态转移,所以增加到8个可能路径。汉明距一栏的第1个值是上一次幸存路径与接收序列的汉明距。第2个值是当前状态转移产生的汉明距的增加。可以看到,在这个例子中总汉明距最小值为2,对应由上一状态 b 转移到当前状态 c。回溯这5个时隙到第1个时隙,已经显示表明时刻1的输入比特是一个0。这一比特从接收端输出,再选出4个新的幸存路径,路径的第1个时隙就丢弃了,然后下一时隙再重复刚才的运算(图中并未显示)。

表7.3 维特比算法,5个时隙的操作,$K=3$,码率为1/2

	可能的路径							
间隔 →	1	2	3	4	5	前一	当前	汉明距
接收比特 →	01	01	00	11	10	状态	状态	
	11	01	01	11	00	a	a	$2+1=3^*$
	11	01	10	01	11	c	a	$3+1=4$
	11	01	01	11	11	a	b	$2+1=3^*$
	11	01	10	01	00	c	b	$3+1=4$
	00	00	00	11	10	b	c	$2+0=2^*$
	11	01	10	10	01	d	c	$3+2=5$
	00	00	00	11	01	b	d	$2+2=4$
	11	01	10	10	10	d	d	$3+0=3^*$

卷积码编码器被介绍为一个在数字通信中特别有效率以及对纠错非常有效的方法。所描述的维特比算法是在实行卷积编码数字信号的最大似然译码中大大减小运算量和存储设备的一个具有创造性的方法。卷积编码有多有效?实际卷积码的性能有多好?它们的性能是怎样依赖于约束长度 K,码率 $1/v$ 和函数生成式的?这里所说的性能是指对编码与译码信息都采用卷积编码后的剩余差错概率。

为了讨论卷积码的性能,需要注意,差错是在被选择接收的路径与发送的正确路径不相同时产生的。然后,差错概率就可以通过计算一个给定的(不正确)路径偏离特定的正确路径的概率,并且把所有这样的路径的概率取平均得到。因为没有某个路径总是正确的,所以假设所有可能的路径都等可能地发生——事实上,这是维特比算法所实现最大似然译码的基本假设。特别地,为了简单且不失一般性,假设正确路径是全零的路径。这暗示着传输信息是一个全零序列。重新关注这一节中始终采用的 $K=3$,码率为1/2编码器的例子,从图7.10编码器的网格表示法注意到对于这一"正确"的传输比特序列,编码器经历的路径是水平方向最上面的那一条,一直保持在状态 a。如果译码器选择的路径偏离这条水平路径,则就发生一个差错。差错概率由译码器偏离这条全零路径的可能的方式以及发生这种情况的概率来得到。但是需要回顾前面提到的,并且通过网格图7.10清楚见到的,进入与再进入状态的重复性。所以,我们可以从任意时刻开始选择离开这条全零路径,再回到这条全零路径的所有不同的情况。特别地,为了简单起见,选择从时刻0开始。现在列出了离开全零路径,再回到全零路径的所能走的不同的路径,按照与全零路径的汉明距从小到大的顺序排列。注意到第1次离开全零路径是惟一的一条转移到状态 b 并接收比特11的路径。从状态 b 有两条可能的路可转移,其中最短的距离是转移到状态 c 的那一条(注意从状态 b 转移到状态 d 也增加了度量值1,但是它离全零状态就更远了)。从这一状态出发,回到状态 a 的最短距离是直接转移到状态 a 的那条路径。这条路径的汉明距已经为5了。这就是最小汉明距的路径。这一最小的汉明距5称为最小自由距 d_f。可以得到卷积译码器的性能精确地依赖于最小自由距,即最小汉明距。留给

读者自己发现该卷积码译码器比最小汉明距稍大一点的汉明距是 6。符合这一距离的有两条可能的路径：$a-b-c-b-c-a$ 和 $a-b-d-c-a$。同样，留给读者自己发现汉明距为 7 的有 4 条路径。当汉明距增大时，其可能的路径数也增多。

一般地，对于任意卷积码译码器，让函数 $a(d)$ 代表偏离全 0 测试路径，然后再回到该路径的汉明距为 d 的路径个数。偏离正确路径的差错概率 P_e 由下式给出其上界：

$$P_e < \sum_{d=d_F}^{\infty} a(d) P_d \tag{7.33}$$

这里 P_d 代表在接收端有 d 个传输比特发生错误的概率。请再一次注意，d_F 被定义为最小自由距离。所有汉明距必须是大于或等于这个距离的。那么，为什么式(7.33)是上界呢？因为很多路径是相互重叠的，正如刚刚举过的那个 $K=3$、码率等于 1/2 的编码器的例子所示。注意，分枝 $a-b$ 对于所有来自全 0 的路径都是相同的。然而，式(7.33)的右端假定了路径都是相互排斥的，因而这些概率可以累加。式(7.33)表明了最小自由距离 d_F 在确定差错概率中的集中性作用。

然而，这个路径差错概率并不是所想要的性能度量。实际上决定编码器性能的是比特差错概率 P_B，它的衡量相对复杂并且可以找出像式(7.33)所示的一个上界。分析 P_B 的具体细节可参阅 Schwartz(1990)和 Proakis(1995)的著作。关键的一点是两个差错概率——路径差错概率和比特差错概率都依赖于 d_F。对于很多信道模型，都可以得到差错概率。如果足够小的话，这是经常需要的条件，就会随着最小自由距的增加而呈指数衰减。在那种情况下，两种差错概率基本上是同阶的，因此，最小自由距越大，则性能越好。对于不同约束长度与编码效率的卷积码，利用计算机搜索找出最大可能的自由距的码的方法在许多年以前由 J.P.Odenwalder 提出。许多这些码，包括后来计算机搜索发现的其他码，都在 Michelson 和 Levesque(1985)以及 Proakis(1995)的著作中制成了表格。特别地，约束长度为 $K=9$，码率为 1/2，应用在第三代 cdma2000 蜂窝系统中，并且在前面图 7.8 中所表示的卷积码正好就是 Michelson 和 Levesque (1985)在其著作的表 9.1 中所列出的。对最小自由距 d_F 的一个有用的界是由 Heller 得到，并且如下式(7.34)所示，它清楚地表明了最小自由距随着 K 和 v 的增加而增大(Proakis,1995)。

$$d_F \leqslant \min_{l \geqslant 1} \left\lfloor \frac{2^{l-1}}{2^l - 1}(K + l - 1)v \right\rfloor \tag{7.34}$$

(符号 $\min_{l\geqslant 1}$ 表示对于遍历所有大于或者等于 1 的 l 值，计算其后的函数值，选择其中最小的值。符号$\lfloor x \rfloor$表示取小于等于 x 的最大的整数)。因此，增加这些参数中的一个或者全部都能够改善卷积码的性能。但是，我们不希望 K 太大，因为维特比译码算法在每个二进制时隙要进行 2^{K-1} 次比较，并且进行二进制判决要基于前面 $L=4K$ 个时隙。增加码率参数 v 需要增加传输速率。对于本节一直作为例子引用的 $K=3$，码率为 1/2 的卷积码验证一下式(7.34)，发现 d_F 的界为 5，正好就是刚才沿着错误路径找到的最小的汉明距的值。事实上，这个界非常接近用计算机搜索得到的 d_F 的值(Proakis,1995)。

采用卷积编码所得到的实际比特差错概率的降低依赖于所选用的信道传输类型。例如，考虑一个在第 6 章 6.4 节简要讨论的加性高斯噪声信道。在那里，指出应用 PSK 是这种信道下的最佳的调制技术，为了保证 10^{-5} 比特差错概率需要信噪比(SNR)为 9.6 dB。$K=3$，码率为 1/2 卷积码编码器采用 PSK，使得需要的 SNR 降低为 6 dB。增加 K 到 5 和 7，保持码率为 1/2，

则进一步分别降低 SNR 到 5 dB 和 4.2 dB。这是对所需的信号功率的一个相当大的降低。第 6 章表 6.1 进一步指出对于一个衰落信道，当 CDMA 系统采用双分集的 RAKE 和卷积编码时可以得到所需信干比(SIR)的非常显著的降低。如该表所示，单独使用 PSK，并不采用任何分集或者编码，将需要信干比为 14 dB 和 24 dB，分别对应的比特差错概率为 0.01 和 0.001。引入双分集的 RAKE 和卷积编码将使得该值降低到可以忍受的值 6 dB 和 7 dB。这些变化显示出了所需信号功率的相当大的降低，所以，卷积码为蜂窝系统的性能带来了真实的改善。在下一节通过介绍 turbo 码的概念，进而对编码技术的讨论进行总结，turbo 码和卷积码都被一些第三代蜂窝系统所采用。

7.3 Turbo 码

在 7.1 节讨论了用分组码进行检错和纠错。把焦点集中在系统码上，即数据比特先输出，然后输出奇偶校验比特。前面一节总结了卷积码的各种情况，从分析过程可以看出卷积码是非系统码，因为它的输出比特是由输入比特与前面输入延时的模 2 和得到的。读者有可能会问，为什么不结合两种码的优点来提高各自单独使用的性能？两种码的结合或者级联技术事实上首先是由 G.D.Forney 在许多年前提出的(Forney，1966)①，他研究的是两个分组码的串行级联。后来，许多研究者致力于一种广泛应用在深太空通信中的分组码和卷积码的级联方案的研究(Michelson 和 Levesque，1985)，这样的一个串行级联框图如图 7.12 所示。分组码编码器首先进行编码，其编码结果进行块交织，使连续的比特在若干块之间分散分布，从而减小发生突发误码的可能性，然后把结果送入卷积码编码器。通常选用的分组码是一类特定的分组码，称为 Reed-Solomon 码，并通常称之为外码。面相应地，卷积码则称为内码。在接收端，执行相反的过程：接收比特流首先经过卷积码译码器进行译码，然后进行解交织，得到的结果再进行分组码译码。两个码的级联能够提供有效的纠错能力，与非级联码达到同样性能的需求相比，能够实现较小的复杂度，以及约束长度较小的卷积码(Michelson 和 Levesque，1985)。

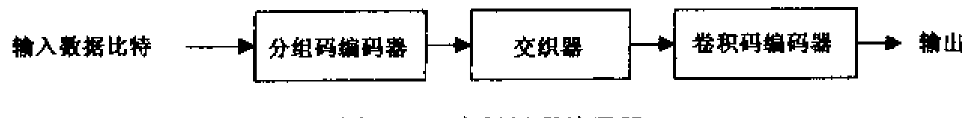


图 7.12 串行级联编码器

我们将看到，本节的主题 turbo 码采用的是两个卷积码的并行级联，并以系统码的形式实现。它们改进了卷积码的性能，这是在高速通信时所需要的，所以这种码目前在许多地方被采用，例如，在第 10 章我们将看到的第三代蜂窝系统，它们代表着纠错码领域的一个相对近代的补充，已经被 C.Berrou，A.Glavieux 和 P.Thitimajshima(1993)②引入通信协会(Berrou 和 Glavieux，1996)③。

它们对这个码的所有性能的介绍引起了编码研究协会极大的兴奋，因为这些码能够提供近似最优的编码性能。其间多年，关于这类码的大量的工作被展开，包括专门针对无线通信的

① Forney，G. D. 1966. *Concatenated Codes*，Cambridge，MA，MIT Press.

② Berrou，C.，A. Glavieux，and P. Thitimajshima，1993. "Near Shannon-limit error-correcting coding and decoding: turbo codes，" IEEE International Conference on Communications，ICC93，Geneva，Switzerland，May，1064 - 1070.

③ Berrou，C. and A. Glavieux. 1996. "Near optimum error correcting coding and decoding: turbo codes，" *IEEE Trans. on Communications*，44，10 (October)，1261 - 1271.

研究。现在有许多书都提供了这方面的介绍与教导。例如，由 Vucetic 和 Yuan 撰写的一本书(Vucetic 和 Yuan, 2000)①。要理解 turbo 码的介绍所带来的兴奋，和它们在无线系统中迅速与其他码的结合应用，必须解释最优编码性能的含义。"最优"一词在这里具有一个特定的意思。它是用香农意义来定义的：Claude Shannon，杰出的工程师和数学家，信息论领域的创始人，在他的 1948 年和 1949 年所著的不朽之著中指明通信信道存在一个容量 C，单位如比特每秒，当它以某一小于容量的速率 R bps 进行传输时，它就有可能采用合适的编码来达到无误。例如，考虑最普通的，也是香农清楚地发现信道容量的例子，就是加性高斯噪声信道或者称为 AWGN 信道：输入待传输的信号进行编码，形成一个连续高斯分布的信号，平均功率为 S，而信道带宽为 W Hz，在信号传输过程中引入了平均功率为 N 的加性高斯噪声。该带宽为 W Hz 的信道的容量由以下这一众所周知的等式来表示(Michelson 和 Levesque, 1985)：

$$C = W\,\mathrm{lb}\left(1 + \frac{S}{N}\right) \tag{7.35}$$

观察香农对信道容量的定义式，式(7.35)是它最常见的书写形式，它并没有指明达到容量的编码方式。许多研究者多年来均致力于找到能够尽可能接近信道容量的码[这一容量被发现可以通过所谓的 M 维正交信号来接近，但是只有随着 $M \to \infty$ 时才可以达到(Schwartz, 1990; Michelson 和 Levesque, 1985)。相干的 M 维 FSK 信号就是这样一个例子]。

继续进一步分析这个例子，考虑小数值信噪比 S/N 下的容量。这些 S/N 的小数值可以通过让带宽 W 无限增大来得到，因为知道噪声功率为 $N = N_0W$，N_0 是噪声功率谱密度(参见在 6.4 节中的讨论)。信道是功率受限的。对于小数值的 S/N，利用近似式 $\ln(1+x) \approx x$，$x \ll 1$ 可以把式(7.35)化成

$$C = WS/(N\ln 2), \quad S/N \ll 1 \tag{7.36}$$

现在有 $S = E_bR$，E_b 是比特信号能量，R 是上面提到的传输比特速率。也注意到 $\ln 2 = 0.69$。于是，留给读者自己推导式(7.36)可以简化为

$$R/C = 0.69/E_b/N_0 \tag{7.37}$$

为了保证传输无误，速率 R 必须比香农容量 C 小，所以由式(7.37)得到

$$E_b/N_0 > 0.69 \tag{7.38}$$

在熟悉的分贝的度量下，这就对应于$(E_b/N_0)|_{\mathrm{dB}} > -1.6$ dB。这一极限被称为香农极限。当 E_b/N_0 大于这一极限时，差错概率理论上可以通过选择一种合适的纠错编码而接近于零。注意，虽然这一极限表达式是由香农在 AWGN 信道下用类高斯的传输信号得到的，但是现在也已经发现这一结果对于其他类型的 AWGN 信道也是可用的。上面所述的 M 维信号就是一个例子，其目的更重要的是展示出在大量衰落信道的类型中也同样存在着一个这样的香农极限。惟一的约束是存在加性白噪声。更细致的讨论可以在 Verdu(2002a)②和 Verdu(2002b)③的著作中看到。

---

① Vucetic, B. and J. Yuan. 2000. *Turbo Codes: Principles and Applications*, Boston, MA, Kluwer Academic Publishers.

② Verdu, S. 2002a. "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, 48, 6(June), 1319-1343.

③ Verdu, S. 2002b. "Recent results on the capacity of wideband channels in the low-power regime," *IEEE Wireless Communications*, 9, 4 (August), 40-45.

现在回到式(7.35)中，重新把这条等式写成 E_b/N_0 的形式，并使用上面提到的关系式 $S=E_bR$ 和 $N=N_0W$。于是，得到下面关于香农容量的等价表达式

$$\frac{C}{W}=\mathrm{lb}\left(1+\frac{R}{W}\frac{E_b}{N_0}\right) \tag{7.39}$$

这一表达式联系了3个参数 C/W(归一化容量)，R/W(归一化传输速率)和 E_b/N_0。特别地，假设非常接近容量，即令 $R=C$，则式(7.39)的香农容量表达式就给出了在任意容量下信号比特能量 E_b 所需的最小值。我们把这个值称为 $E_{b\min}$，重新书写式(7.39)得到以下形式：

$$\frac{C}{W}=\mathrm{lb}\left(1+\frac{C}{W}\frac{E_{b\min}}{N_0}\right) \tag{7.40}$$

这一参数表达式把 C/W 与 $E_{b\min}/N_0$ 联系起来。如果绘制 $E_{b\min}/N_0$ 关于 C/W 的曲线，就得到了香农容量曲线，它给出了在任一要求的容量下，达到零差错概率所需要的最小可能的信号传输能量。这一最小能量随着容量关于带宽归一化值的增大而增大(Michelson 和 Levesque，1985)。在 C/W 很小的特殊情况下，再次得到了式(7.38)所示的极限 $E_{b\min}/N_0=0.69$。

如何把所熟悉的信号与式(7.40)的极限曲线进行比较呢？在前一节曾指出，当PSK信号传输经过了一个加性高斯信道，也就是正在讨论的信道时，对于比特差错率为 10^{-5} 所要求的信噪比 SNR = 9.6 dB。然而，需要注意的是，第5章所讨论的奈奎斯特信号成形指出了PSK的比特速率 R 与传输带宽 W 的比值接近于1。令 $R=W$，也就是假设 $S/N=E_b/N_0=9.6$ dB。为了与香农最佳结果比较，让式(7.40)中的 C/W 为1[注意在书写式(7.40)时，令 $R=C$]，则发现 $E_{b\min}/N_0=1$ dB 或者 0 dB。PSK 在比特差错概率为 10^{-5} 下需要比零差错概率的绝对最小值多出 9.6 dB 的信号功率。注意，在前一节还指出，对于 $K=7$，码率为1/2的卷积码，相同的信道需要 4.2 dB，比不编码的 PSK 改善了 5.4 dB，但仍然与最佳值 0 dB 之间存在着 4.2 dB 的差距。而在 Berrou 等(1993)和 Berrou 与 Glavieux(1996)的著作中，所讨论的 turbo 码通过仿真发现它在 10^{-5} 比特差错概率下需要的 $E_b/N_0=0.7$ dB。这使得这些码与香农最佳值的差距在 0.7 dB 以内(为了仿真需要，认为 10^{-5} 是接近于0的)，所以，它非常令人兴奋！

刚才脱离正题讨论了最优码，现在回到正题，开始简要地介绍 turbo 码。turbo 码包括了所谓的递归系统卷积码(RSC，recursive systematic convolutional)的并行级联，通过一个交织器将两个这样的 RSC 码分开。图7.13展示了一个码率为1/2的递归系统卷积码的例子(Berrou 和 Glavieux，1996)。"系统"一词的意思已经陈述过，就是指输入信息比特直接送到输出。图7.13的输出线 c_2 可以作为这个用途来使用，这条线上的第 i 个比特正好是第 i 个输入比特 d_i。"递归"意味着卷积码编码器的输出是反馈到输入端的，这同样可以从图7.13中看到。拿一个码率为1/3的 turbo 编码器作为例子，它就有像图7.14的形式(Berrou 和 Glavieux，1996；Vucetic 和 Yuan，2000)。这两个递归系统卷积编码器 RSC1 和 RSC2 并行地运转，提供了3个输出比特中的2个比特。注意，输出数据比特是直接送到上一个卷积编码器中的；它们通过一个交织器再送到下一个编码器。输入比特也同时直接送到输出端作为第3个输出比特，正如一个系统码所要求的那样，具体见图。交织器的设计对 turbo 码的性能非常关键，在此，性能是以比特差错概率来定义的(Berrou 等，1993；Berrou 和 Glavieux，1996；Vucetic 和 Yuan，2000)。交织器的目的是为了使输入两个编码器的比特序列相关，也是为了使接收端的译码器重新得到输入的比特序列。伪随机扰码，使用很长的数据块长度，被发现能够提供最佳的交织器性能：交织长度越长，

性能越好(Vucetic 和 Yuan,2000)。turbo 码的性能还依赖于汉明式的自由距(Berrou 和 Glavieux,1996;Vucetic 和 Yuan,2000),所以,我们可以通过一个对汉明距的搜索过程找到好码。

一个在讨论卷积码译码时提到的与此过程类似的最大似然方法用来决定接收端译码器的形式。结果译码器呈现出和发射端相同的由交织器分隔的两个译码器的串行级联(Berrou 等,1993;Berrou 和 Glavieux,1996;Vucetic 和 Yuan,2000)。为了得到最佳性能需要进行连续迭代。对于一个码率为 1/3,如图 7.14 所示的 turbo 码译码器的例子,如图 7.15 所示(Vucetic 和 Yuan,2000)。符号 c_0^*,c_1^* 和 c_2^* 分别用来表示图 7.14 中发送的 3 个比特信号在接收端所对应的接收信号。发送信号假设已经经过了一个已知属性的信道,例如,具有随机衰落,或者加性高斯噪声等。图 7.15 所示的串行连接的两个最大似然译码器 ML1 和 ML2 的特定形式依赖于信道的属性,以及图 7.14 的编码器 RSC1 和 RSC2 的形式。注意到,接收信号 c_0^*,c_1^* 送到第 1 个译码器 ML1;接收信号 c_2^* 送到第 2 个译码器 ML2。同时,ML1 的输出的交织结果,以及接收信号 c_0^* 的交织结果也送到 ML2 作为输入。ML2 的输出再反馈给 ML1(迭代),得到第 3 次迭代的输入。就是这样一个反馈的回路使得这一码系统得到"turbo"码的命名,它是基于 turbo 引擎的基本原则的。需要注意,ML 译码器、交织器和解交织器都引入延时,这一点在图 7.15 中并没有显示出来。迭代将继续进行,一直收敛到一个需要的最佳可能性能时才停止。对于大的交织长度,例如 4096 比特或者更长,为了达到收敛,需要迭代次数的典型值为 12 次到 18 次(Vucetic 和 Yuan,2000)。

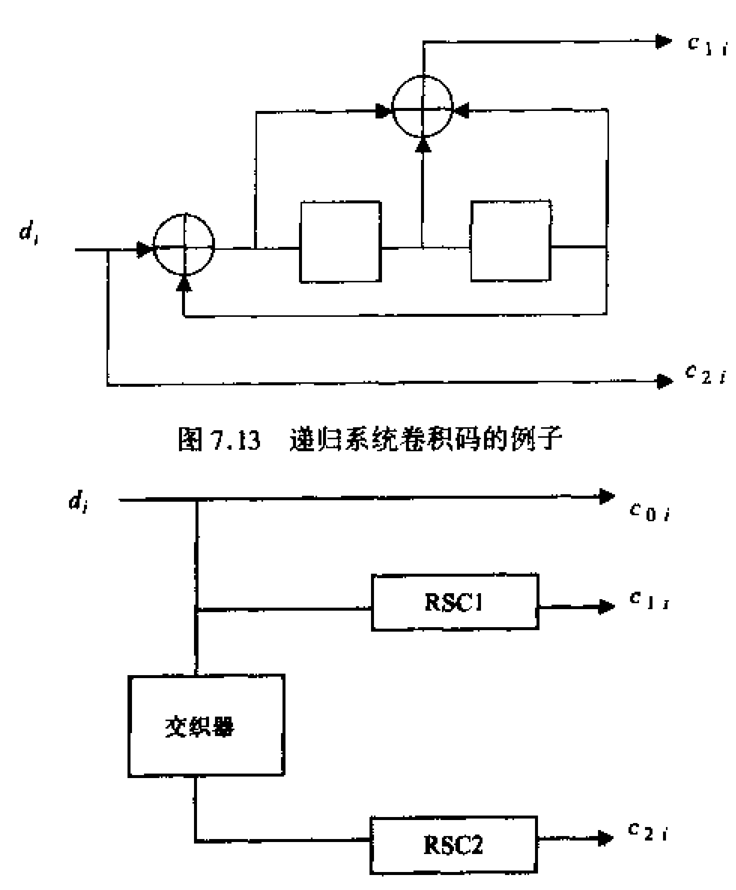


图 7.13 递归系统卷积码的例子

图 7.14 码率为 1/3 的 turbo 编码器

在瑞利衰落的信道下使用 turbo 码得到的效果已经由 Hall 和 Wilson(1998)提出,并在 Vucetic 和 Yuan(2000)的著作中进行了总结。对 1/3 码率的 turbo 码的仿真结果表明,在这一环

境下，它们能够获得与这一信道可能的容量相差 0.7 dB 以内的性能。这显示了该码相对于未编码的 PSK 具有超过 40 dB 的编码增益！（回顾第 6 章衰落信道上 PSK 为了达到比特差错概率 0.001 需要的 $E_b/N_0 = 24$ dB。在比特差错概率为 10^{-5} 时需要的信噪比增加到 44 dB）。Vucetic 和 Yuan(2000)在著作中还包含了一些 turbo 码应用在第 10 章所讨论的第三代蜂窝系统中的信息。

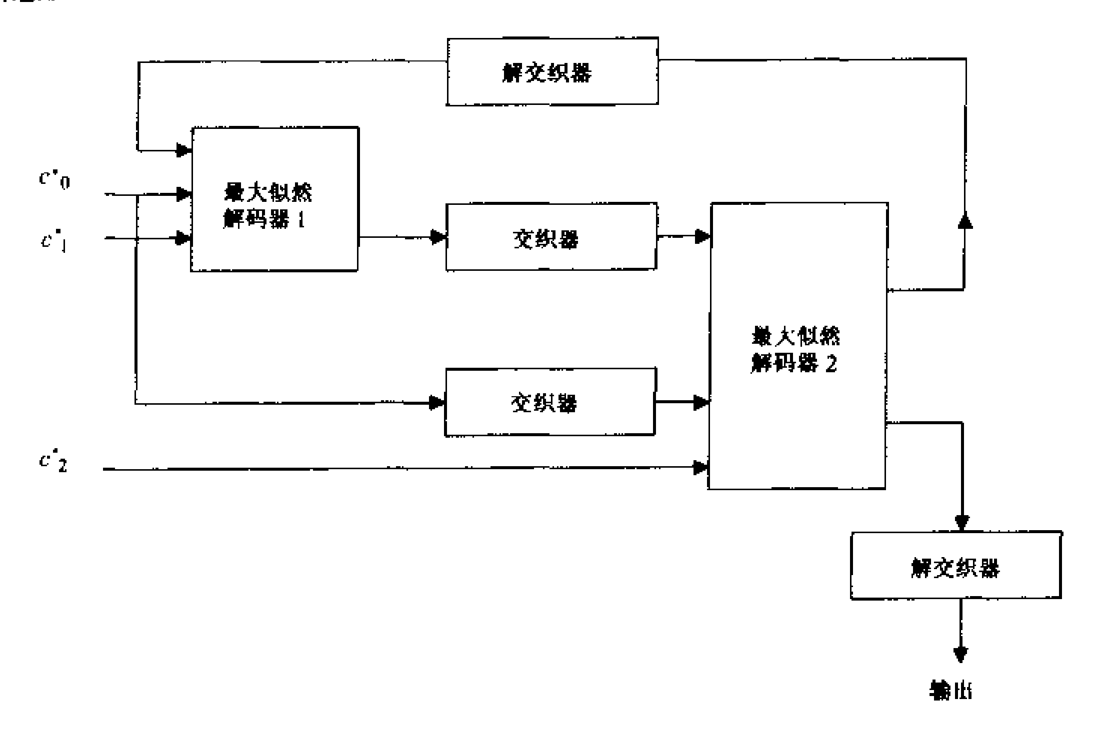


图 7.15 图 7.14 的 turbo 编码器的迭代的译码器

习题

7.1 (a) 说明 1 个奇偶校验比特，令它等于 k 个信息比特组成的数据块的模 2 和，能够用来检测信息比特中一个偶数的差错，并用一些不同长度的数据比特来检验这一结果。

(b) 对于奇校验，重复(a)的要求。

7.2 考虑式(7.2)的(7,3)码

(a) 说明它能够生成如表 7.1 所示的 8 个码字；

(b) 说明这个码可以用式(7.6)、式(7.7)和式(7.15)分别对应的 $\boldsymbol{P},\boldsymbol{G},\boldsymbol{H}$ 矩阵来表示；

(c) 写出这个码的 $\boldsymbol{H}^T$ 矩阵；

(d) 写出当接收码字向量第 2 位出错时，表 7.1 的每一个码字所对应的伴随式 $\boldsymbol{s}$。把这些向量与 $\boldsymbol{H}$ 矩阵的合适的列进行比较。

7.3 系统分组码(7,4)码具有如下的生成矩阵 $\boldsymbol{G}$

$$\boldsymbol{G}=\begin{bmatrix}1&0&0&0&1&0&1\\0&1&0&0&1&1&1\\0&0&1&0&1&1&0\\0&0&0&1&0&1&1\end{bmatrix}$$

(a) 找出这个码的 $\boldsymbol{P}$ 矩阵，并与(7,3)码的 $\boldsymbol{P}$ 矩阵进行比较；

(b) 找出监督矩阵 $\boldsymbol{H}$ 和它的转置$\boldsymbol{H}^T$,并与(7,3)码进行比较;

(c) 这个码有多少个码字? 把它们全部列出来。这个码可以纠正多少个错误? 这个是否与汉明界相符?

(d) 为什么(7,3)码和(7,4)码被认为是对偶码?

(e) 为什么这个码具有 8 个可纠正的错误图案,包括零错误向量在内? 找出每个错误图案的伴随式。

7.4 用式(7.19)的汉明界来说明(15,11)码能够纠正 1 个差错;(10,4),(11,5)和(15,8)能够纠正两个差错;(15,5),(23,12)和(24,12)能够纠正 3 个差错。分别计算这些码的编码效率并做比较。

7.5 (a) 检验式(7.23)对 x^7+1 因式分解的 3 个因式是正确的。

(b) 找出(7,4)码的两种可能的形式,分别计算它们的生成矩阵 $\boldsymbol{G}$,并说明它们之中有一个与上面习题 7.3 中的码是一致的。

(c) 用式(7.24)来检验表 7.2 所列出的(7,3)码的 8 个码字的生成。

7.6 (a) (7,3)码的生成多项式列在表 7.2 中。按照式(7.28)计算该码的每个非零输入数据序列的余式多项式 $r(x)$,从而找出它们各自的奇偶校验比特。

(b) 检验图 7.3 所示的(7,3)码编码器确实能够生成该码合适的奇偶校验比特。

7.7 用图 7.2,找出习题 7.3 中(7,4)码的编码器,类似(7,3)码的编码器图 7.3 那样。找出一些输入数据序列的奇偶校验比特,并写出相应的 7 比特码字。检验该码能够纠正单比特差错。

7.8 考虑 3 个 cdma2000 的帧质量指示器移位寄存器的编码器实现,如图 7.4 到图 7.6 所示。它们生成多项式分别为式(7.29)到式(7.31)。

(a) 在 $p=10^{-5}$和 10^{-4}两种情况下,计算当数据块分别为 250,500 和 1000 比特时未能检测的差错概率各是多少,并对结果做出讨论评价。

(b) 考虑帧质量指示器为 8 比特的图 7.6 或者式(7.31)。选择任意一个输入 10 比特长的模式并说明当它离开编码器时输出为 18 比特的系统码。

7.9 考虑图 7.7 的 $K=3$,码率为 1/2 的卷积码编码器。

(a) 初始化编码器使得输入只有 0。初始化后输入如下比特序列,并跟踪输出比特序列

1 1 0 0 1 0 1 1

(b) 检验图 7.9 的状态图能够表示该编码器。当输入序列为(a)题所示,编码器状态从 00 开始,跟踪编码器的状态转移。指明输出序列与(a)题的结果一致。指明图 7.10 和图 7.11 的网格图和树状图都能够表示该编码器。

7.10 (a) 对图 7.8 所示的 $K=9$,码率为 1/2 的卷积码编码器,输入任意比特序列,跟踪器输出比特序列。注意,编码器首先必须初始化。

(b) 检验它的两个关联向量能够用十进制形式的 753 和 561 来表示。

7.11 考虑一个 $K=3$,码率为 1/3 的卷积码编码器,生成式为 $g_1=[1\ 1\ 1]$,$g_2=[1\ 0\ 1]$和 $g_3=[0\ 1\ 1]$。

(a) 画出该编码器框图。

(b) 像图 7.9 那样,给出该码的状态图,标注不同的状态,以及状态之间的转移。找出当输入为习题 7.9(a)所说的比特序列时输出比特序列是什么。注意,编码器的初始化。

7.12 (a) 检查表 7.3 的维特比算法。确认经过 5 个时隙后得到的 8 个幸存路径,以及当前时刻的 4 个新的幸存路径是否如表所示。选择一个任意较长时隙的接收比特序列,重复该算法,每个时隙对输入的前 5 个时隙做出判决。

(b) 选择一个至少有 15 比特的任意输入序列。它们用图 7.7 的 $K = 3$,码率为 1/2 的编码器进行编码,跟(a)题一样。写出编码输出序列,然后对于下面两种情况进行维特比译码:首先是无误的情况,其次是接收序列在传输过程中出现了一些差错。比较两种情况下的估计结果。首先用判决深度 $L = 5$,然后尝试 $L = 12$。对比两种情况下的结果。

7.13 对习题 7.11 中的卷积码编码器重复 7.12(b)。列出类似表 7.3 的表格。写出在无误以及若干误码情况下的结果估计。尝试两种译码深度 5 和 12,并比较结果。

7.14 (a) 关于卷积码编码器的路径的汉明距的讨论,对于图 7.7 所示的编码器,以及图 7.10 所示的网格图表示法,指明其最小汉明距,即最小自由距是 5。并指出汉明距为 6 的两条路径和汉明距为 7 的 4 条路径。

(b) 检验式(7.34)的上界对于(a)题中的编码器给出了最小自由距为 5。习题 7.10 和习题 7.11 中的这个上界值是多少? 这个界是如何因为约束长度和码率的不同而改变的?

7.15 (a) 按照 7.3 节的分析,说明为什么香农极限当 $C \ll W$ 时是 -1.6 dB。这里 C 是信道容量,W 是传输带宽。这个值表现出通信的带宽非常宽。

(b) 现在考虑限带的通信,例如无线移动通信。这里 W 不像(a)中的那么大。在这个例子中说明香农极限最佳值是 0 dB。提示:正如课本中所说,R/W 接近于 1,在这里选用合适的信号成形,R 是在该信道中的传输速率。同时要注意,香农理论指出的是有可能达到 R/C 接近于 1,并且通过合适的编码达到零差错概率。

第8章　第二代数字无线系统帧

本章讨论当前的第二代数字蜂窝系统。第6章在介绍TDMA,CDMA以及对它们的容量进行比较时,就已经涉及了这些系统的部分内容。特别讨论了基于TDMA的北美IS-136(D-AMPS)系统,它是继承第一代AMPS模拟蜂窝系统的数字系统;基于TDMA的欧洲GSM系统,除欧洲外,它也应用于美洲和世界其他地方;基于CDMA的IS-95系统。下面对这些系统进行详细的研究,首先描述通过移动台和基站之间的空中接口建立和接收呼叫的信号。然后,讨论连接基站,控制从一个蜂窝到另一个蜂窝移动切换的有线网络内部的信号,以及用于定位和寻呼移动台以便传送呼叫的信号。在本章的最后,将简单讨论用于这些系统的语音信号处理和编码技术。

第二代蜂窝系统的简要框图如图8.1所示。基站(BS)控制了多个移动台(MS)[北美标准将移动台称为蜂窝用户站(CSS)]。移动交换中心(MSC)轮流控制多个基站,负责处理蜂窝之间的切换、移动台定位、寻呼以及其他移动管理和控制功能。本地位置寄存器(HLR)包括所有移动用户的参考信息和侧面信息,作为在MSC注册的"本地位置"。异地访问者在本地的MSC注册。异地访问者的参考信息和侧面信息在与移动台的HLR通信后,存储于与本地MSC有关的访问位置寄存器VLR中。

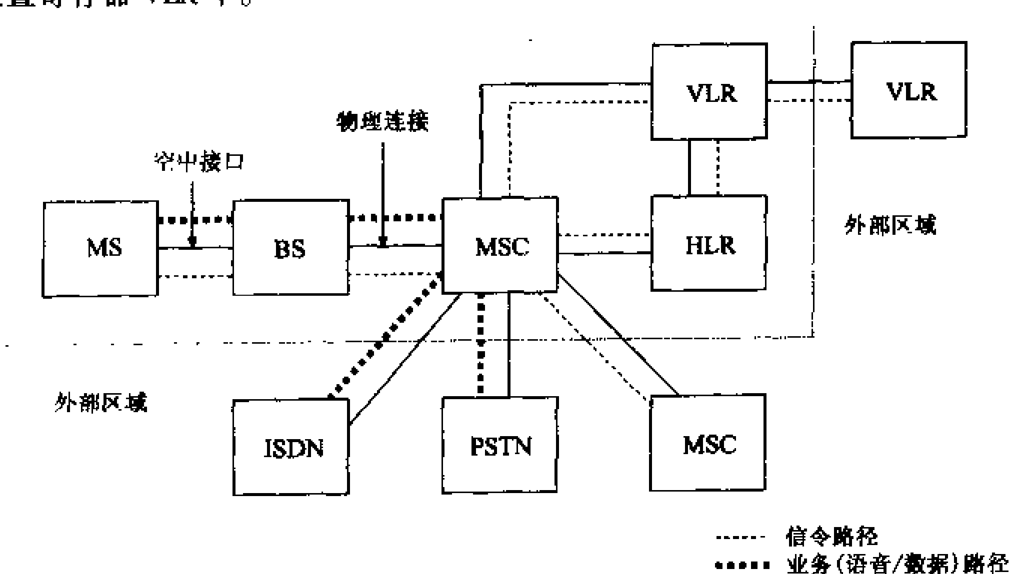


图8.1　简化的蜂窝网

启动移动台的注册与认证,不论是预先发送信号或接收信号,都是通过移动台和基站之间的空中接口发送适当的控制信息实现的。然后,这些信息被送往MSC进行认证。如果移动台位于本地位置,MSC询问它的HLR,核实和批准该移动台的注册信息。如果移动台位于异地位置,本地MSC/VLR联合将移动台的注册信息传送给它的HLR。图8.1中的细虚线表示不同控制信号的传输路径。粗虚线表示在呼叫建立和进行中语音和数据流的传输路径。通信经过公共电话交换网(PSTN)或综合业务服务网(ISDN)完成。

8.1 节将讨论 GSM 系统中不同的控制信号经过空中接口用于移动台注册和建立连接的问题。8.2 节将讨论 IS-136/D-AMPS 系统中相应的控制信号问题。8.3 节将讨论基于 CDMA 的 IS-95 系统。8.4 节描述用于这些系统中的切换、定位、寻呼过程。最后一节简要描述语音信号处理和编码技术。经过空中接口传输前完成数字语音流和所有控制信号的差错控制编码，以便减小传输过程中衰落、噪声、干扰的影响。

8.1 GSM

GSM(Global System for Mobile Telecommunication)上行链路(移动台到基站)工作在 890 MHz 到 915 MHz 频段，下行链路(基站到移动台)工作在 935 MHz 到 960 MHz 频段。每个链路 25 MHz 带宽被划分成多个 200 kHz 的信道，每个频段在低频点留有 200 kHz 的保护带宽。在每个传输方向上有 124 个 200 kHz 的信道可用。每个信道使用 8 个时隙的重复帧结构传输信息，如第 6 章所述(参见图 6.3)。全速率的信道每一帧占用一个时隙。如果网络运营商需要，移动台和基站每一帧可以从一个信道跳转到另一个信道，以减小衰落的影响。这里假设衰落是频率选择性衰落；在一个信道上经历衰落的移动台可以通过跳转到其他信道来改善传输。为了避免干扰，这种做法需要每一个蜂窝和邻近的蜂窝进行相当多的协调工作。

4.615 ms 长的 8 时隙帧和放大后的比特结构如图 8.2 所示。时隙一共有 148 个比特，保护时间 30.5 μs，相当于连续帧中有 8.25 个分离比特(控制信号时隙在时间上具有相同的时隙长度，但含有不同的字段)。注意到图 8.2 中所示的时隙起始和结束于 3 比特全零同步字段。在时隙中间位置的 26 比特训练序列用于估计无线电信道，训练接收端自适应均衡器，帮助克服可能遇到的多径衰落。该时隙具有 144 数据比特，分成两组，每组 57 比特。两个标志比特 F 指示数据字段载有用户信息或是控制信息(后面将会看到，用户数据信息可能会被特殊的控制信息打断)。系统在 0.577 ms 时隙间隔传输 156.25 比特，正如前文所述，比特率为 270.833 Mbps。用户数据速率为 22.8 Kbps：每 4.615 ms 长的帧传输 114 比特数据，但是每 13 帧中有 1 帧预留给控制信号。

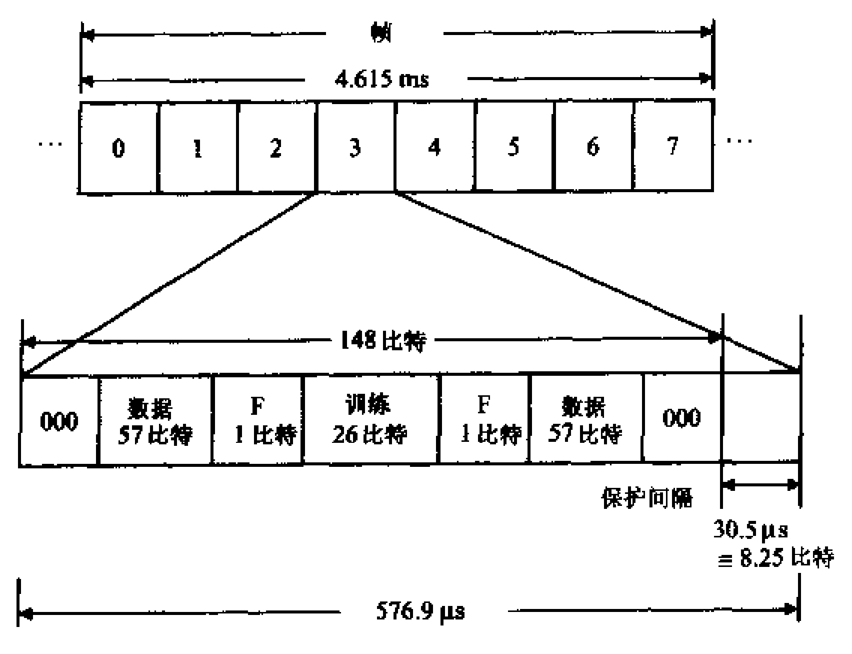


图 8.2　GSM 的帧/时隙结构，业务信道时隙

简要总结完 GSM 用户时隙后，下面考虑建立呼叫的过程。图 8.3 中描述了很多控制信息，它们将经过空中接口发送出去，这些信息需要特殊的 GSM 帧和时隙。在下面各节中将进行简要介绍。

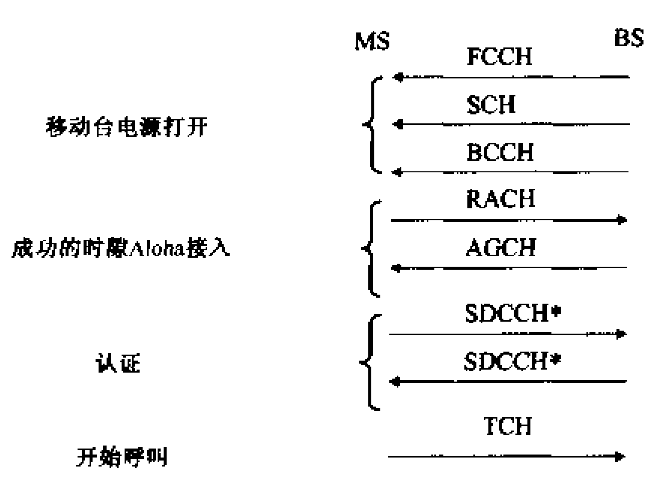


图 8.3 一个 GSM 呼叫的建立①

很明显，移动台在建立呼叫或准备接收呼叫前，需要打开电源。移动台在任意一个蜂窝中打开电源，必须首先跟踪获取蜂窝中使用的频率，然后进行帧时隙结构同步。为了实现同步，首先寻找特殊的控制信道——频率校正信道(FCCH)，这些信道被本地基站广播，用于移动台调整频率特性，实现与基站的同步。FCCH 信息通常跟随同步信道(SCH)信息用于辨认基站和提供帧，所以与移动台同步(注意，这里的信道和信息可以互用，控制信道通常载有指示信息)。一旦频率或时间信息需要，移动台监听广播控制信道(BCCH)，该信道提供建立呼叫所需的信息：蜂窝配置、所属网络、接入信息、控制信道信息和其他一些信息(Steele，1992：Ch.8②；Goodman，1997③)。

利用这些信息，移动终端可以发起一次呼叫。首先，向随机接入信道(RACH)发送随机接入请求信息，该信息载有 5 比特随机数和 3 比特用途指示(用途包括尝试新呼叫、回复寻呼信息、位置注册及需要传送的短信息)。在后文将要提到的时隙 Aloha 随机接入协议，用于发送和接收尝试接入的确认信息。因为同一时刻可能有多个终端随机尝试接入基站，从而导致碰撞，所以称为随机接入尝试。在发生碰撞的情况下，基站无法正确接收 RACH 信息，因而无法发送确认信息。时隙 Aloha 协议规定，传输站点(移动台)在没有接收到确认信息的情况下，在等待一个随机的时间间隔后重新尝试接入。如果该接入尝试成功，基站利用接入准许信道(AGCH)信息确认接收 RACH 信息，如图 8.3 所示。该信息重复 8 比特请求信息，指示终端连接到单独专用控制信道(SDCCH)，移动台通过该信道传输认证所需的信号信息，建立所需的呼叫连接。如果该呼叫被批准，基站用 SDCCH 信息回复移动台，指示移动台连接到频率/时隙通信信道(TCH)，从而开始实际的呼叫并发送用户所需要传送的信息(多数情况下是数字语音信息)。

这些控制信道实际上是如何工作的呢？这些信道加上下面将要提到的 3 个其他控制信道，作为逻辑信道出现在图 8.2 所示的帧/时隙结构中。这些控制信道被分成 3 类：广播信道、通用控制信道和专用控制信道。在每一组中定义了不同信道，如表 8.1 所示。表中所示的

---

① 有可能是包含多条消息的一个消息系列。

② Steele, R. ed. 1992. *Mobile Radio Communications*, London, Pentech Press; New York, IEEE Press.

③ Goodman, D. J. 1997. *Wireless Personal Communication Systems*, Reading, MA, Addison-Wesley.

3 个广播信道，前面已提到它们用于建立呼叫，参见图 8.3，如同寻呼和接入控制信道一样，都是定向下行链路，即从基站到移动台。寻呼信道用于呼叫到来时基站定位移动台。图 8.3 所示的随机接入信道是定向上行链路，即从移动台到基站。表 8.1 所示的专用控制信道是双向的，允许移动台信号、管理信息、监督信息双向传输。特别是基站使用的下行链路慢速随路控制信道（SACCH），向移动台发送发射功率等级和定时预备指示。移动台反过来利用该信道上行链路向基站发送接收信号长度、通信控制信道质量、邻近蜂窝提供的用于移动协助切换的广播信道测量指示信息。快速随路控制信道 FACCH 用于发送切换请求和其他紧急信号信息。它被发送到普通的业务信道（TCH）或 SDCCH 中，并中断该信道。为此目的请回顾图 8.2 中标志比特 F 的使用。

表 8.1　GSM 控制信道列表

广播信道（所有下行，BS→MS 方向）
1. 频率校正信道，FCCH
2. 同步信道，SCH
3. 广播控制信道，BCCH
通用控制信道
1. 寻呼信道，PCH（下行，BS→MS）
2. 接入准许信道，AGCH（下行，BS→MS）
3. 随机接入信道，RACH（上行，MS→BS）
专用控制信道（双向）
1. 单独专用控制信道，SDCCH
2. 慢速随路控制信道，SACCH
3. 快速随路控制信道，FACCH

现在考虑这些不同逻辑信道在物理时隙和帧结构中的实际工作情况。移动台和基站通过每帧包括特殊时隙的通信信道（TCH）相互发送数据信息，如图 8.2 所示，采用分配好的频率对。在整个呼叫过程中，该频率对可以固定不变，也可以按照规定的跳跃规则在每帧间隔发生变化。控制信道按照同样的方式工作，利用给定的频率信道分配规定的时隙控制信道。单独专用控制信道（SDCCH）用在从移动台到基站指定专用双向连接 TCH 信道之前，如图 8.3 所示，它使用的是指定数量的频率信道中的专用时隙。GSM 利用 3 种不同的方式为其他控制信道分配时隙：设置规定数量的频率信道，每个方向有 124 个可以被利用，每 1 帧预留 1 个时隙用于控制；每 13 帧分配 1 个时隙，组成通常的业务信道时隙序列，作为控制信道；如果需要的话，通过设置 TCH 标志比特，使用规则的业务信道时隙。我们将简要讨论表 8.1 中所示的不同控制信道的实现方法，细节请参见 Rappaport（1996）[①]，Goodman（1997）和 Steele（1992：Ch.8）的著作。

每 13 帧分配一个时隙按照业务信道序列组成控制信道，是通过定义 120 ms 长的重复 26 帧复帧结构实现的，如图 8.4 所示。业务信道占用 0～11 帧和 13～24 帧。第 12 帧被标记为 S，如图 8.4 所示，被指定为表 8.1 中专用慢速随路控制信道（SACCH）。如果需要的话，第 25 帧也可以这样指定。反之，它就是一个空闲帧，被标记为 I。S 帧中的每一个时隙与相应的业务信道相关联，为该信道提供所需的信号信息。SACCH 时隙利用同样的比特结构作为业务信道，如图 8.2 所示。由于每一个这样的时隙载有 114 个信息比特，SACCH 中每 26 帧占用 1 个时隙，因此比特速率为 114 比特/120 ms，即 950 bps。如果每个复用帧有两个时隙，则比特速率就加倍，

① Rappaport, T. S. 1996. *Wireless Communications, Principles and Practice*, Upper Saddle River, NJ, Prentice-Hall PTR.

即 1900 bps。这就是该信道被指定为慢速随路控制信道的原因(注意到用户的业务速率 24 倍于此速率,为已经提到的 22.8 Kbps。半速率业务信道也可以定义,在 26 个复帧中分配 12 个时隙,而不是全部的 24 个时隙,数据速率为 11.4 Kbps。在这些半速率的 TCH 信道中每个复帧分配一个 SACCH 时隙)。我们将会看到 SACCH 信息包含 456 比特长的编码信号信息,该信息在 4 个复用帧时间间隔里进行交织,占用 480 ms。如果时间间隔对于移动管理太长,例如在切换的时候,通过设置 F 比特中断 TCH 信道,快速随路控制信道(FACCH)信息通过该信道发送。

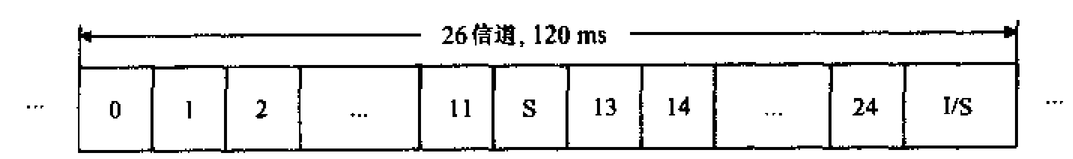


图 8.4　重复 26 信道的多帧结构,GSM

前面已经看到有 3 种方法用于 GSM 系统控制信道分配,除了 SDCCH 以外,它利用指定频率信道中的专用时隙。前面描述的两种方法,每 13 帧分配一个时隙给 SACCH 并在 TCH 中设置 F 比特,提供 FACCH 信息,包含表 8.1 中两个专用控制信道。第 3 种方法是将数量可选的频率信道中的 0 时隙取出,用于实现表 8.1 中剩余非专用控制信道。在上行链路方向,利用这些选择的频率,所有帧中的 0 时隙被分配给随机接入的 RACH 信道。表 8.1 中剩余的广播和通用控制信道都是定向下行链路,从基站到移动台都被分配了 0 时隙,每一帧都是重复的 51 帧复用帧控制结构,使用下行链路选择频率集合(当然,其他的时隙被分配到 TCH 或 SDCCH 信道)。该 51 帧复用帧结构如图 8.5 所示。在该图中,这 51 帧每 235 ms 重复一次,被分成 5 组,每组 10 帧,每一组中的第 1 个信道,在图 8.5 中标记为 F,与频率校正信道(FCCH)相对应(回忆图 8.3,这是移动台用于锁定基站频率的信道)。每一个 FCCH 紧随同步信道(SCH),SCH 用于移动台建立帧、时间和同步。该信道在图 8.5 中被标记为 S(注意到,单词 immediate 指的是查阅下一帧的 0 时隙,所有不同帧的参考只查阅这些帧的时隙 0)。在第 1 组中,SCH 帧后面的 4 帧载有广播控制信道(BCCH),在图 8.5 中被标记为 B,该信道用于提供建立呼叫所需的信息。BCCH 每 51 帧只出现 4 次。图 8.5 中其余信道被标记为 C,与接入信道(ACH)或寻呼信道(PCH)相对应。寻呼信道用于在呼叫到来时在特定蜂窝中定位移动台。寻呼将在 8.4 节中进行讨论。

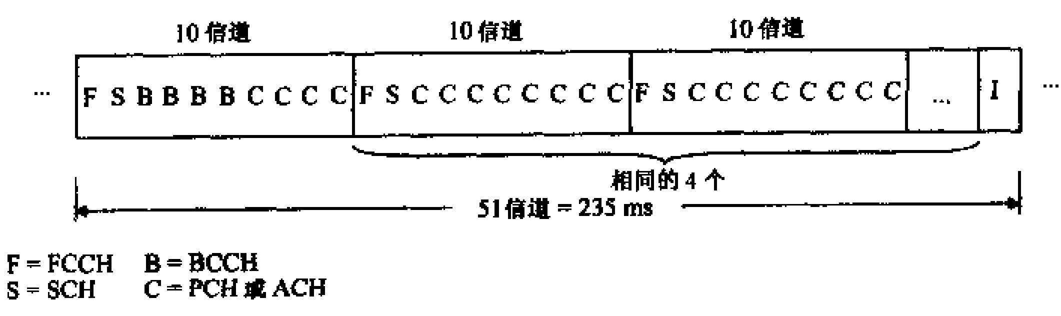


图 8.5　GSM 下行多帧控制结构:被选择的频率信道的第 0 时隙

如图 8.2 所示,GSM 时隙都是 576.9 μs。业务信道时隙包含 148 比特,30.5 μs 保护间隔。除了随机接入信道(RACH),所有控制信道每个时隙都包含 148 比特。除了 FCCH 和 SCH,其余所有信道使用的是和 TCH 相同的时隙比特分配,如图 8.2 所示。频率校正信道(FCCH)传输 148 比特全零序列。这导致在移动台用于排列其自身所生成到基站的频率的全部时隙间隔

内,GSM基站可用来传输非调制正弦信号。基站传输的SCH载有比TCH更长的64比特训练序列,所以它载有较短的78比特数据有效负载,如图8.6(a)所示。图8.6(b)显示了如何从25比特信息中产生78比特数据字段。该信息被10比特CRC和效率为1/2的卷积码保护。这两种差错控制方法在第7章中已经讨论过。CRC提供外编码,卷积码执行内编码。第7章中讨论的卷积码用于连续的输入数据。这里讨论的卷积码以及后文讨论的卷积码,用于有限长度的数据块保护。在每个数据块结束时,卷积码必须被适当地终止,从本质上说,就是当每一个连续的数据块结束时,在相同的状态下重新开始卷积码的编码操作,尾比特就用于这种目的。在图8.6(b)这种特殊的情况下,4个尾比特被添加到数据块的尾部。在第7章中,曾提到在码率为1/2的卷积码中为每一个信息比特发射两个编码比特。带有10个CRC比特和4个尾比特的25比特信息,在图8.6(b)中变成78比特数据字段。

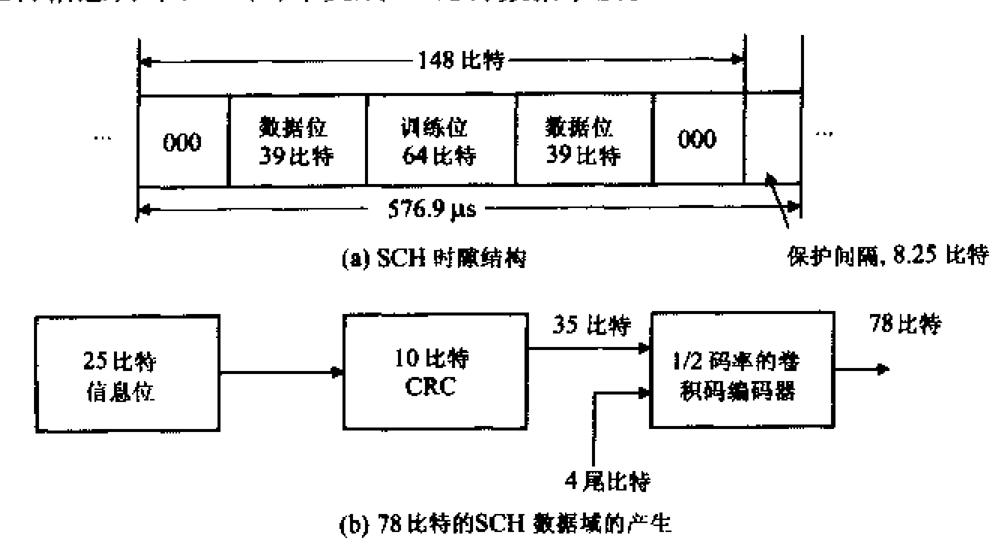


图8.6 SCH实现,GSM

原来的25比特信息包含6比特基站身份码和15比特帧同步序列(Steele,1992:Ch.8)。建立GSM呼叫(参见图8.3)时,移动台和基站之间的同步通过将移动台锁定到一个15比特序列来实现,每一个序列被编码为78比特数据(参见图8.6和图8.5)。

考虑表8.1中除RACH以外的其余控制信道[包括广播控制信道(BCCH)、寻呼信道(PCH)和接入准许信道(AGCH)]。每一个控制信道时隙的114比特数据字段(参见图8.2)是通过将每一个复用帧连续4个时隙456比特进行交织而产生的。这456个比特通过184个比特信息产生,该信息被效率为1/2的卷积码加上由40个校验比特构成的差错校正编码所保护。4个尾比特被加在224比特编码序列后面进入卷积码编码器。图8.7描述了456比特的产生过程。由图8.5可以看出,字母B表示的广播控制信道,每51个信道复用帧出现一次,使用连续4帧的0时隙传输经过编码的184比特信息。图8.5中字母C表示的寻呼和接入允许控制信道,将编码后的184比特信息适当地组成4个连续帧。

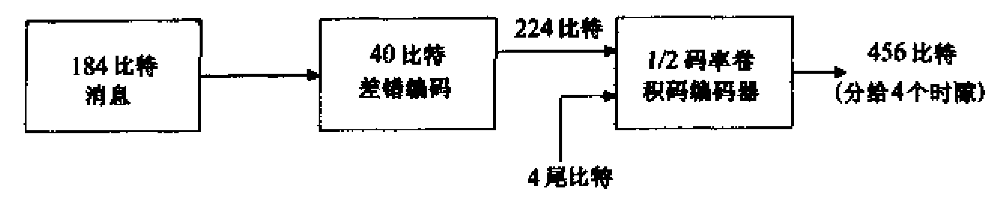


图8.7 每时隙数据域的114比特的生成,BCCH、PCH、AGCH:GSM

与其他信道不同，随机接入信道(RACH)的时隙比特分配结构如图8.8所示。实际的突发传输较短，只有88比特。保护间隔相应为68.25比特，即252 μs。既然移动终端初始时可能在蜂窝中的任何一个位置，那么RACH信息的保护间隔作为移动台，通过上行链路发送给基站，并且需要回复的第1个信息必须为蜂窝中最大可能的迂回路由传播延迟留有足够的时间。选择保护间隔使得移动台可以与最大距离为35 km(Steele,1992:Ch.8)的所在蜂窝的基站通信(注意，传播速度为300 000 km/s，对于此距离的传播延时为233.3 μs，低于保护时间的长度)。如图8.8(a)所示，8个和3个尾比特分别定义了突发数据流的开始和结束，41个比特用于同步。36个比特数据字段中的其余比特将原始的8比特随机接入信息，包括5比特随机数加上3比特目的指示进行编码，如图8.3所示。图8.8(b)显示了6个奇偶校验比特加在8比特信息后面提供差错校正度量，然后将合成的14比特编码信息加上4个尾比特进行卷积编码，提供36比特的RACH数据字段。

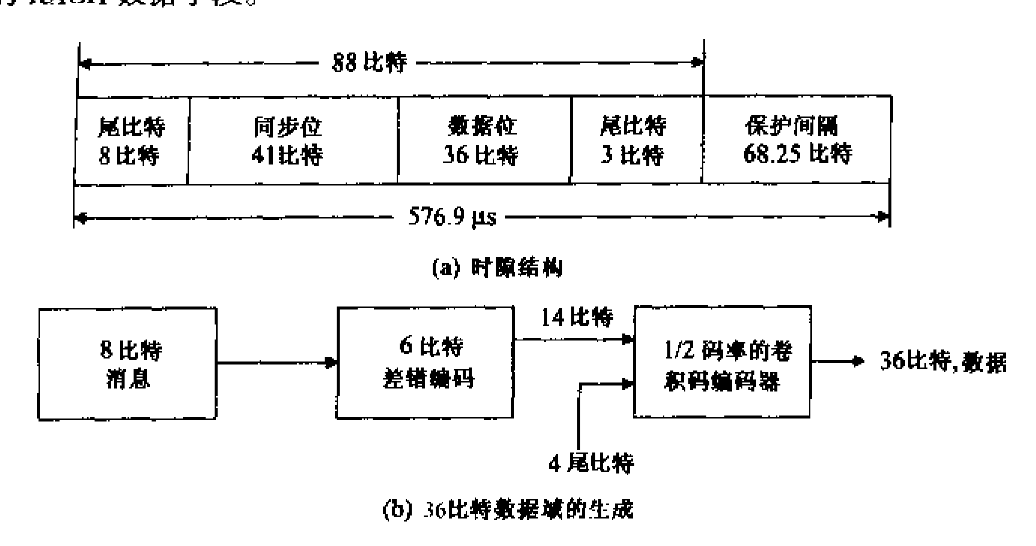


图8.8 随机接入信道，RACH：GSM

8.2 IS-136或D-AMPS

本节讨论第6章中提到的其他基于TDMA的第二代主要数字蜂窝系统IS-136或D-AMPS，以及其业务信道和建立与维护呼叫所需的逻辑控制信道。该第二代系统出现在GSM系统之后，实际上可以发现其控制信道和GSM系统中的控制信道有一些相似之处。

从前几章的讨论中可以看出，该系统在北美使用，占用25 MHz频段，上行链路从824 MHz到849 MHz，下行链路从869 MHz到894 MHz。在这些频段内，每30 kHz划分一个频率信道，每个频率信道包含重复的TDMA帧，每帧载有6个时隙。与GSM系统中的TCH相比，数字业务信道时隙结构的例子如图8.9所示。如前所述，每帧中有两个时隙被分配给全速率用户。

系统传输速率为48.6 Kbps：每40 ms传输1944比特/帧。全速率用户每帧传输520数据比特，虽然发送速率为48.6 Kbps，但是数据传输速率为13 Kbps。这些数值与GSM系统有很大不同。前面提到，D-AMPS系统与以前模拟AMPS系统使用相同的频率信道结构，所以在低数字利用率的情况下，双模手机可以很容易地在模拟系统中使用。

从图8.9中可以看出，上行和下行两个方向上的时隙结构不同。上行方向需要6比特保护时间G，因为蜂窝中的移动台可能相对于基站移动，导致与基站的距离变化，这将阻止终端

同时开始通信,防止相互干扰。在上行链路相同时隙结构中的电源跳动时间 R 用于给可能没有启用的终端供电,这些终端需要时间使传输功率跳到所需数值,该数值依赖于所在小区的位置和基站测量到的干扰功率。由于基站(BS)[①] 总是处于工作和传输状态,下行链路时隙结构不需要这两个字段。下行链路,从基站到移动台,时隙载有 12 比特 DL 字段。该字段载有 7 比特载波位置信息,指示数字控制信道的载波频率。该信息被 4 个奇偶校验比特保护。12 比特中有一个被保留。

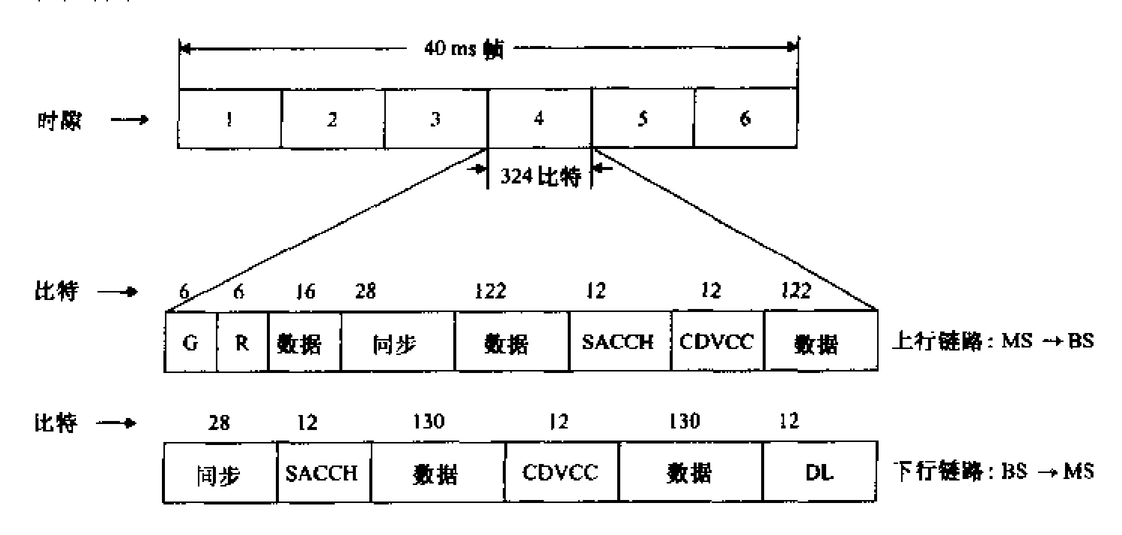


图 8.9　IS-136 时隙结构,数字业务信道

12 比特 CDVCC(coded digital verification color code)字段包含 8 比特 DVCC 数,再加上 4 个用于保护的奇偶校验比特。该字段用于连续握手:基站传输该数值;移动台用相同的数值进行回复。如果没有回复或回复不正确,则该时隙被丢弃。每一个方向上的 28 比特同步字段用于建立时间同步和训练接收均衡器。每一帧中的 6 个时隙同步字段具有不同的同步序列。每一个方向上的 12 比特 SACCH 字段载有慢速随路控制信道。和 GSM 系统中的 SACCH 相比,该信道用于在移动台和基站之间发送控制信息。例如,移动台回复基站发送的测量顺序信息,进行接收信号强度和比特差错率的测量,以及其他定量测量,而且向基站发送测量结果。这些用于决定切换到另一个蜂窝是否正当(这个在移动台辅助基站的测量过程被称为移动台辅助的切换,或者 MAHO 过程)。全速率业务信道,每帧包含两个时隙,每帧有 24 比特用于 SACCH。在这种情况下,SACCH 控制数据传输速率为 600 bps。在 GSM 系统中,快速随路控制信道 FACCH 通过将业务信道中每个时隙的 260 比特替代为控制信息来获取。对于每帧两个时隙的全速信道而言,该信道控制数据传输速率为 13 Kbps。

该系统是如何建立和维护移动台与基站之间的通信的呢?逻辑控制信道 DCCH 用于这个目的。全速 DCCH 与每帧两个 48.6 Kbps 时隙相对应,在可选择的频率信道中通常是时隙 1 和时隙 4。需要建立通信的移动台打开电源,寻找到 DCCH。该 DCCH 中的信息使它定位指定的 DCCH,然后移动台使用 Aloha 接入过程,就像 GSM 系统中一样,注册、监听任何寻呼,再开始呼叫。IS-136 系统中逻辑控制信道,包含 DCCH,如表 8.2 所示。通过后面的讨论,读者可以将它们与 GSM 系统进行比较。这里定义了 3 种下行逻辑控制信道。下行广播控制信道 BCCH,包

① IS-136 协议使用了术语基站、*MSC*(*移动交换中心*)和交互作用函数,即 BMI,来代替基站。这里为简单起见,将其简写为 BS。

含快速广播控制信道(F-BCCH),用于承载对时间敏感的信息,该信息用于移动台建立分配到的 DCCH 结构,以及注册和接入信息;扩展广播控制信道(E-BCCH)载有对时间不太敏感的信息;SMS 广播控制信道(S-BCCH)用于控制 IS-136 中定义的广播短信息服务(GSM 也有类似的短信息服务。这些服务允许短数据信息和通常的语音信息在这些蜂窝中传输。数据传输容量将在后面的章节中结合第三代分组交换系统和基于数据的无线局域网进行详细的讨论。这里集中讨论只有语音通信的第二代系统)。

表 8.2 IS-136 的 DCCH 逻辑控制信道列表

下行(BS→MS)
1. 广播控制信道,BCCH
快速广播控制信道,F-BCCH
扩展广播控制信道,E-BCCH
SMS 广播控制信道,S-BCCH
2. SMS 点到点、寻呼和接入回复信道,SPACH
寻呼信道,PCH
接入响应信道,ARCH
SMS 信道,SMSCH
3. 共享控制反馈信道,SCF
上行(MS→BS)
随机接入信道,RACH

SMS 点到点、寻呼和接入回复信道(SPACH)是一个逻辑信道,用于承载寻呼和接入回复控制信息,以及关于 SMS 服务的点到点信息。共享控制反馈信道(SCF)用于承载从基站到移动台的下行链路的信息,作为随机接入过程的一部分。该信息用于指示一个随机接入时隙给移动台,同时提供随机接入请求的确认。用于呼叫建立或启动的随机接入信息,以及移动台注册和认证,使用了一个上行链路,随机接入信道(RACH),这和 GSM 系统一样。

这些不同的逻辑控制信道是如何实现的?首先,集中讨论下行广播信道。如前所述,这些信道在全速情况下,每帧使用两个时隙进行传输,在选择性频率信道中通常是时隙 1 和时隙 4(我们只关注全速情况)。广播信道使用重复超帧结构持续传输。超帧结构如图 8.10 所示。每一个超帧与 16 个连续的 TDMA 帧相对应,占用时间间隔 640 ms。每一个 6 时隙 TDMA 帧由两个 3 时隙块组成。每块中有一个时隙(如上所示,通常是时隙 1 和时隙 4)分配给 DCCH,所以每个超帧有 32 块长。如图 8.10 所示,3 ~ 10 个块分配给 F-BCCH 逻辑信道。后面依次跟随 E-BCCH,包含 1 ~ 8 个块,S-BCCH 分配 0 ~ 15 个块,如果需要的话,0 ~ 7 个块预留给其他功能,最后剩下的 32 块超帧分配给 SPACH。然而,SPACH 至少被分配一个块。很明显,它的最大长度为 28 块(为什么?)。

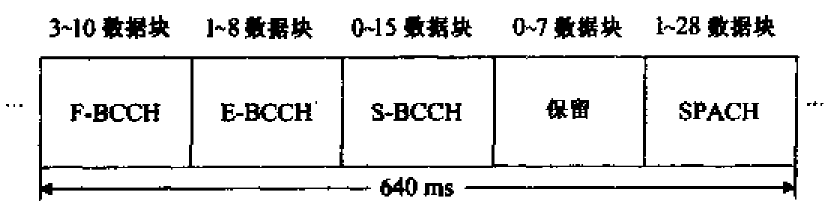


图 8.10 IS-136 的重复超帧结构

信道中的 F-BCCH 信息具体指定块是如何分配给每一个逻辑信道的?打开电源的移动台,寻找到一个 DCCH,指示它连接到指定的 DCCH。该 DCCH 中的 F-BCCH 载有包含 DCCH 结

构的信息：超帧中 F-BCCH、E-BCCH、S-BCCH、保留、SPACH 块的数量；接入信息，例如移动接入功率、接入突发大小、随机接入算法的参数（最大重传数、授权的随机数字、SMS 信息长度等）；注册和系统 ID 参数（网络类型、国家代码等）。

所有的下行广播逻辑信道使用 324 比特时隙结构，与图 8.9 中的下行数字业务信道类似。这些控制信道的时隙结构如图 8.11 所示。注意，每一个时隙载有相同的 28 比特同步字段和总长为 260 比特的数据，和数字业务信道一样。然而，控制时隙结构有些地方不同，CSFP 字段替代 CDVCC，22 比特 SCF 替代 SACCH 和 DL 字段。当然，SCF 比特用于承载表 8.2 中的 SCF 逻辑信道。所以，SCF 信息作为控制信道时隙的一部分传输，而非像其他下行链路广播信道信息那样作为控制时隙（眼光敏锐的读者在图 8.10 的超帧格式中，可能会注意并想到 SCF 信道是怎样实现的）。如前所述，SCF 信道用于回复移动台的随机接入尝试。后面会进一步提到该信道载有的信息。继续讨论控制信道时隙结构，我们注意到 12 比特 CSFP（编码的超帧相位）字段用于指示超帧中 TDMA 块的位置。该字段载有 8 比特 SFP（超帧相位）数字和用于保护该数字不出差错的 4 位奇偶校验比特。8 比特中的 5 比特在 32 块超帧中用于保持以 32 为模计数，其余 3 比特被预留而且被置为 0（再次注意我们这里只考虑全速率 DCCH，对于半速率 DCCH，只有偶数位上的 SFP 数字用于计数）。在第 1 个 F-BCCH 时隙，该数字从 0 开始，超帧中的每一个 TDMA 块增加 1。

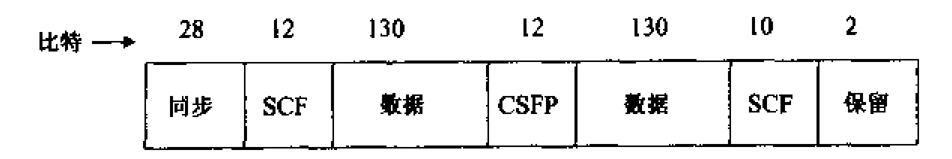


图 8.11　IS-136 时隙结构，下行广播控制信道

下面考虑下行链路 DCCH 时隙结构中的 260 比特数据。它们由 125 比特加上 5 个尾比特产生，将 130 比特组成效率为 1/2 的卷积码编码器，与 GSM 系统中相应的过程类似，如图 8.6(b)和图 8.7 所示。输出 260 位编码比特，通过连续的 TDMA 帧进行交织，进一步防止衰落。输入给卷积码编码器的 125 比特从包含头比特和信息比特的 109 比特产生，后而加上 16 位 CRC 比特进行保护。与表示 GSM 系统相应过程的图 8.6(b)和图 8.7 相比，差错控制编码过程如图 8.12 所示。图 8.13 显示了信道控制信息是如何出现在效率为 1/2 的卷积码编码器的输入端的（对于熟悉网络分层结构的读者来说，IS-136 规则显示出图 8.13 中的信息字段代表第 3 层数据，而在图 8.11 中，对 TDMA 控制时隙中的数据字段，所有 130 比特代表第 2 层数据）。

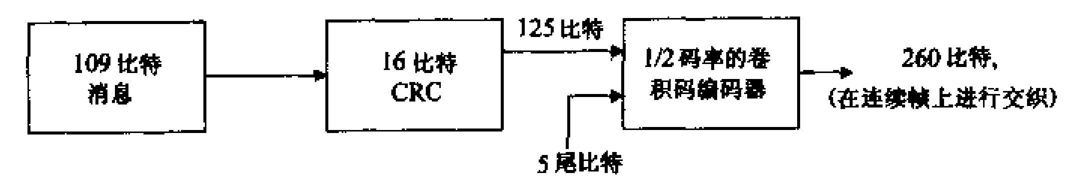


图 8.12　IS-136 控制数据域生成

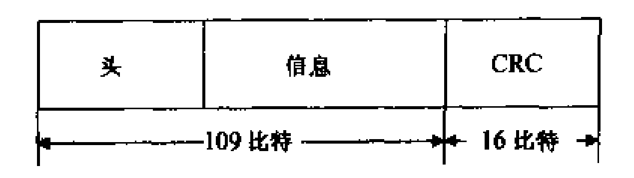


图 8.13　IS-136 DCCH 下行消息的一般形式

将 F-BCCH 信息作为一个例子。该控制信道用于传输 DCCH 结构、接入参数、注册和系统 ID 信息。很多不同类型的信息被发送，包括这个信息。这些信息可能多于 109 比特传输，所以连续信息占用了下一帧。图 8.14 给出了 3 种不同类型的 F-BCCH 帧的例子。如图 8.12 和图 8.13 所示，注意到所有帧都包含 16 比特 CRC 字段、首部信息和长度可变的数据字段，标记为 L3 数据。如前所述，这是第 3 层数据字段，载有一个不同的 F-BCCH 信息。8 比特参数 L3LI 为长度指示器，用 8 进制指示紧随其后的 L3 数据字段的长度。图 8.14(a)代表了(b)和(c)在一帧中当 F-BCCH 信息可以被立即传输时所用的帧结构。如果信息完整，则传输需要多于一帧，图 8.14 中的开始和连续类型将被使用。1 比特的 EI 标志置为 1，表示全 0 被用于拉长帧；EI = 0 表示一个新信息的开始。EC 用于指示 E-BCCH 的变化。帧起始于 1 比特指示信息 0 或 1。图 8.14(c)中连续帧的 7 比特 CLI，即连续长度指示器，用于指示连续信息的比特数。

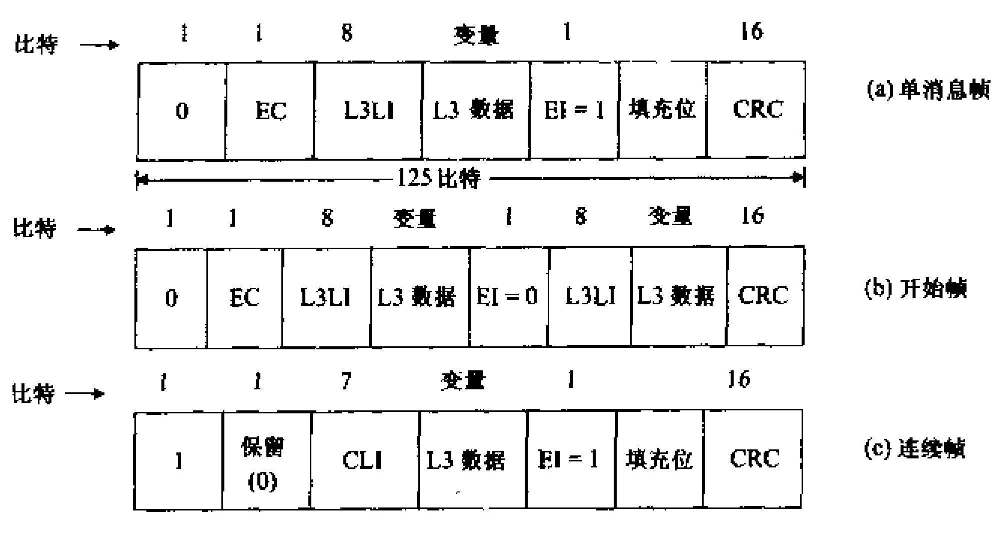


图 8.14 F-BCCH 帧的例子

作为 F-BCCH 信息的一个例子，DCCH 结构信息可以包含 43 比特到 201 比特。接入参数信息为 65 比特长，注册信息长度在 16 比特到 77 比特范围内变化，系统 ID 信息在 27 到($191 + N \times 17$)比特范围内变化。其他 F-BCCH 信息也被定义，有些是命令的，有些是可选的。

下面考虑一个 IS-136 中的上行链路控制信道 RACH，如表 8.2 所示。如前所述，移动台利用该信道开始呼叫、注册、回复寻呼信息。它的功能与前面讨论的 GSM 系统明显类似。如同上面讨论的其他(下行链路)DCCH，在选择性频率信道中，帧中时隙 1 和时隙 4 载有该信道。IS-136 接入过程使用 RACH，与下行信道协同作用，如图 8.15 所示。该过程使用时隙 Aloha 接入协议，在讨论 GSM 接入过程中已经提到过，如图 8.3 所示。移动台监听下行 SCF 信道(如图 8.11 所示)，决定发送 RACH 信息的具体时隙，然后发送 RACH 信息，如图 8.15 所示。随后特定时隙的 SCF 信息，将指示 RACH 信息是否被正确接收，以及接入是否被准许。如果接入被准许，随后的 ARCH 信息将出现在 SPACH 信道(参见表 8.2 和图 8.10)，指示具体的数字业务信道给移动台通信使用。如果接入没有成功(其他移动台可能同时也在尝试接入)，延迟一个随机时间以后再进行接入尝试。

两个时隙结构用于发送 RACH 信息：正常的时隙结构(324 比特长)如前面所述的在 IS-136 中的业务信道和控制信道(参见图 8.9 和图 8.11)，以及缩短的时隙结构。两种结构分别如图 8.16(a)和图 8.16(b)所示。注意到它们的数据段长度不同。正常结构一共载有 244 个数据

比特;缩短结构则载有 200 比特。后者载有相应的 38 比特保护时间(AG)加上 6 比特跳变时间 R。头字段 G,R 和同步字段被保护和跳变,同步字段出现在上行业务信道,如图 8.9 所示。PREAM 为前置放大器字段,包含 DPSK 调制策略中连续 $-\pi/4$ 相位变化,用于帮助符号同步。附加同步字段(sync +)是必须的,因为 RACH 没有被连续使用,如同数字业务信道一样,所以,附加同步时间是必须的。

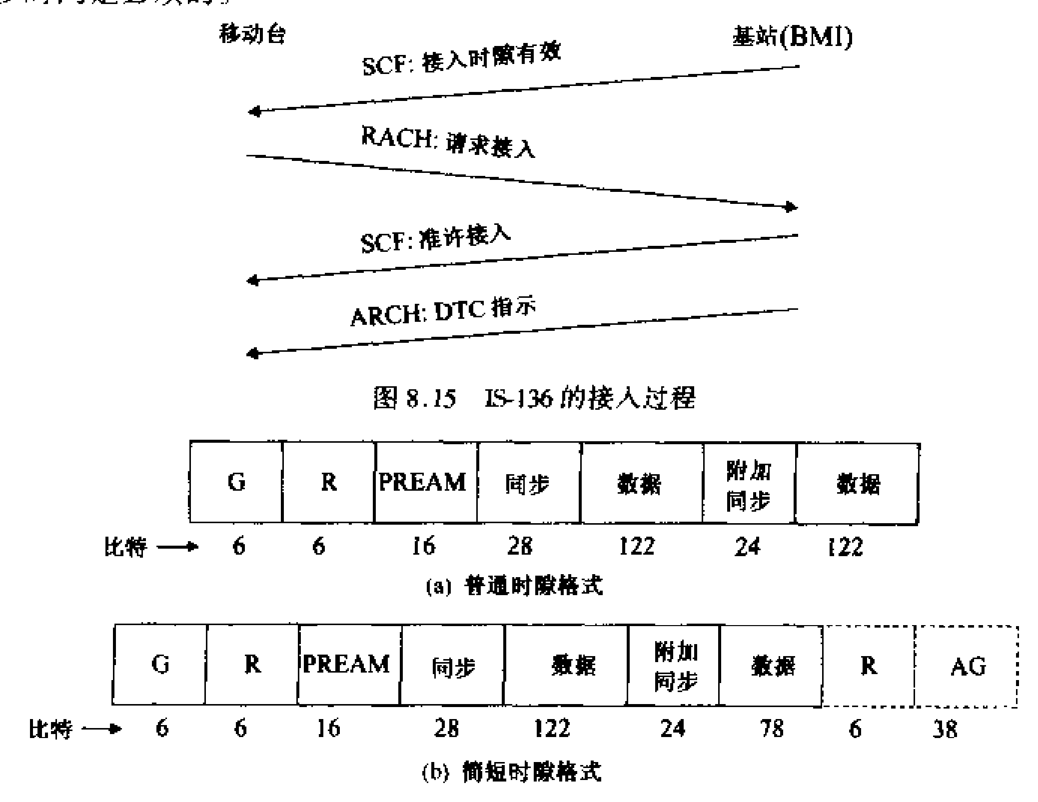


图 8.15　IS-136 的接入过程

图 8.16　IS-136 RACH 的时隙格式

RACH 数据字段的产生方法与其他 DCCH 协议中的数据字段类似。例如,正常的 RACH 信息开始是 101 比特长。16 位 CRC 比特附加在后面用于保护这些比特,然后输出的 117 比特加上 5 个尾比特,再通过码率为 1/2,也用于其他信道的卷积码编码器来进一步保护,如图 8.12 所示。起始的缩短 RACH 信息为 79 比特长,然后利用相同的过程组成 200 比特数据字段大小。一共有 7 种不同类型的 117 比特长的正常 RACH 帧可能被传输。这些包括 BEGIN,CONTINUE,END 和单纯的 BEGIN/END 帧,以及其他帧(与前面所讨论的不同的 F-BCCH 帧相比较,见图 8.14)。例如,BEGIN 帧结构如图 8.17 中所示。3 比特 BT 字段用于区分这 7 种不同类型的 RACH 帧;例如,000 指示 BEGIN 帧。移动台 ID(MSID)有 3 种可能的数值:20 比特临时 MSID;34 比特 MSID,它与通常的 10 比特北美电话号码转换为二进制格式相对应;50 比特国际 MSID。2 比特 IDT 字段指示哪种 MSID 出现在 MSID 字段。1 比特 TH 表示扩展首部的存在:0 表示没有扩展首部;1 表示存在扩展首部。3 比特 NL3M 字段指示有多少与第 3 层相关联的信息出现在 BEGIN 帧中。L3LI 字段的功能与在F-BCCH 结构中相同,如图 8.14 所示。

它指示实际信息字段与后面的 L3 数据字段的长度,单位为字节。8 种不同的 RACH 信息被定义。该信息已经被暗含地提到,认证、寻呼回复、呼叫建立和注册包含在其中。通过包含在信息格式中的 6 比特信息类型字段区分这些信息。起始信息包含呼叫组数,作为它的格式之一。呼叫组数在其中可选。

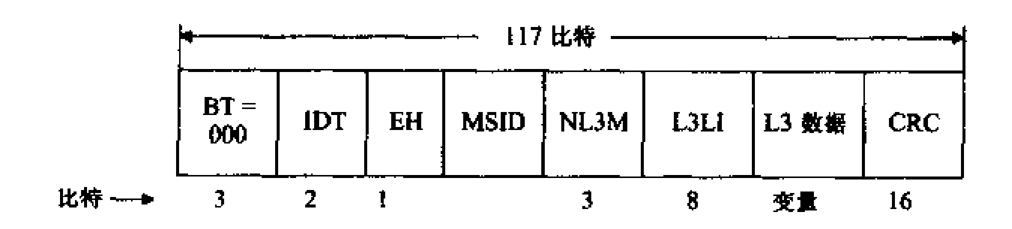


图 8.17 IS-136 RACH 消息的 BEGIN 帧格式

下面转到下行共享控制反馈 SCF 信道,前面在解释接入过程时简要讨论过,如图 8.15 所示。这里在前面简要讨论的基础上进行扩展。由图 8.11 可以看出,DCCH 中载有 22 比特 SCF 信道。如图 8.15 所示,该信道用于通知准备通过给定的 DCCH 开始接入尝试的移动台 RACH 时隙的位置,该时隙用于接入尝试的具体时刻(一次接入尝试可以回复给 SPACH 寻呼子信道(PCH)所接收的一次寻呼;它可能用于注册目的;它可能是一次呼叫建立尝试;它可能用于发送其他 RACH 信息,就像前面提到的那样)。一旦 RACH 信息被特定的 RACH 时隙发送,随后的 SCF 信息,如图 8.15 所示,将承载着信息,以指示接入尝试是否成功。SCF 信道承载的 3 个分离的标志用于指示当前特定的 RACH 时隙是否空闲,可以用于接入尝试;其次,可利用前面指定的 RACH 时隙指示接入尝试是否成功。第 1 个标志称为 BRI 标志,包含 22 个 SCF 比特中的 6 位,用于指示特定的 RACH 状态为忙(111100)、保留(001111)或者空闲(000000)。空闲的 BRI 标志指示 RACH 时隙,在特定的时间间隔之前用于接入尝试。如果前面一次的接入尝试被基站成功地识别,则 BRI 标志被置为忙。5 比特长的第 2 个标志,R/N 标志,用于指示接入尝试是否成功:11111 指示尝试被接收;00000 指示没有被接收。最后的 7 比特被 4 位奇偶校验比特保护,一共 11 比特被指定为 CPE(coded partial echo)标志。当移动台实现了一次成功接入时,这 7 个比特是 MSID 中最不重要的 7 个比特,如图 8.17 所示。注意,3 个标志都用于指示一次成功的随机接入尝试。

8.3 IS-95

让我们回忆前面的讨论,特别是在第 6 章中,IS-95 是一个基于 CDMA 的系统。它的业务信道和控制信道被定义为特定的编码,而不是 GSM 和 IS-136 系统中的时隙。这里的讨论将与前面几节基于 TDMA 蜂窝系统的描述有很大的不同。

首先,讨论业务信道系统块的图表,如图 8.18 和图 8.19 所示。这些图表是对图 6.5(a)中 CDMA 块图表的相当大的扩展,在前面第 6 章已经讨论过,其目的是描述信号扩频的过程,使用尽可能简单的伪随机扩频码。而在这里试图描述得更准确,特别是因为使用码字,而非时隙作为信道来承载通信和控制信息。下面简要描述可用做比较的控制信道系统图表。注意到图 6.5(a)中描述的 CDMA 信号处理的基本部分同样也出现在图 8.18 和图 8.19 中:二进制信息流与伪噪声码片扩频码序列相乘,输出的结果通过适当的低通滤波器成形,然后提供给高频发射器。然而,这里使用模 2 和运算替代图 6.5 中的乘法运算。读者可以自己推导二进制 1 和 0 的模 2 和运算,它与对应的双极性二进制数字的乘法运算(提示:$(+1)\cdot(+1)=+1$;$(-1)\cdot(-1)=+1$;$(-1)\cdot(+1)=-1$;$(+1)\cdot(-1)=-1$)是一致的。图 8.18 显示反向链路即上行链路发射器采用 OQPSK 调制模式(参见第 5 章),图 8.19 中前向链路即下行链路发射器采用 QPSK 调制模式。

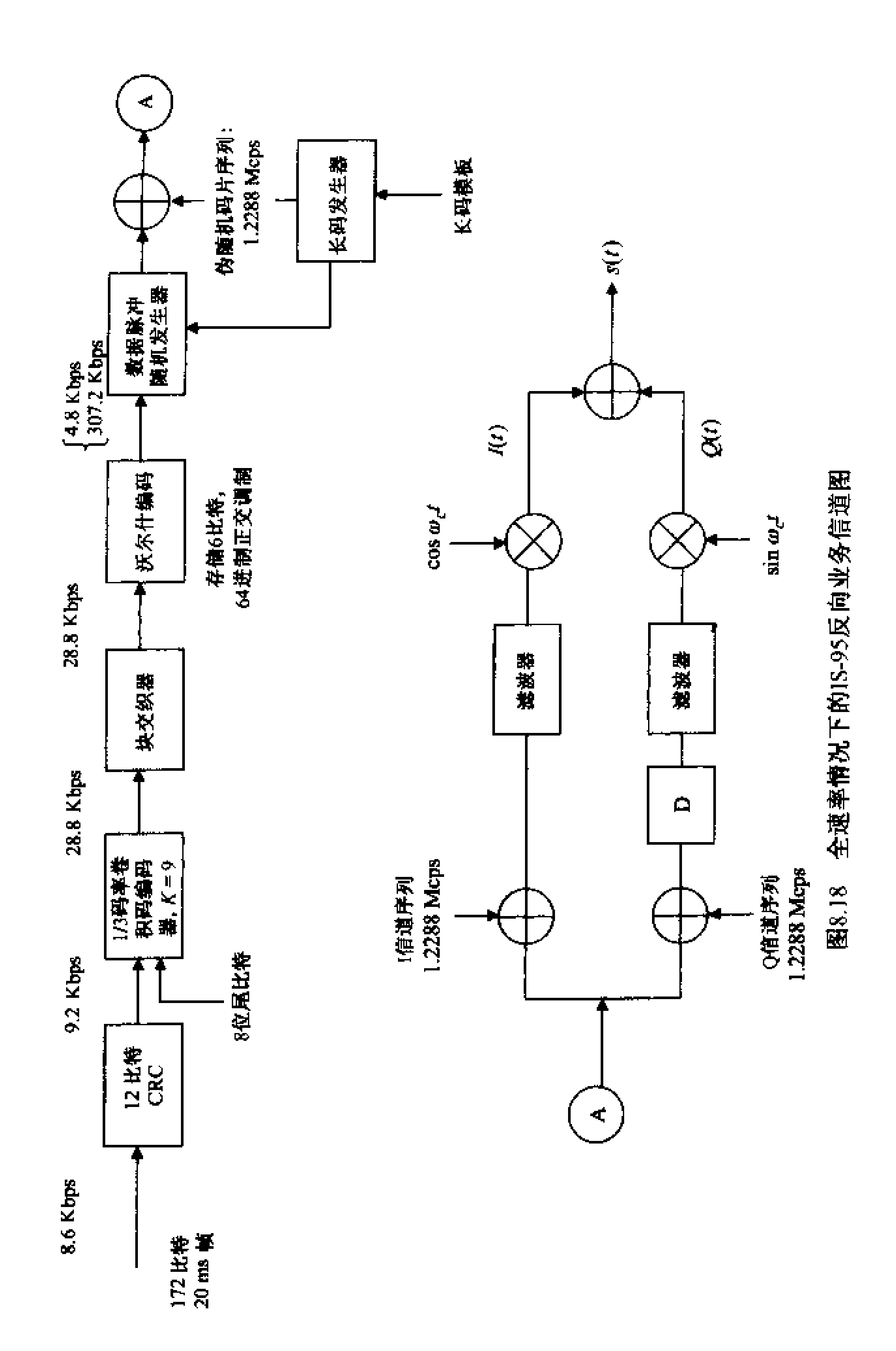


图8.18　全速率情况下的IS-95反向业务信道图

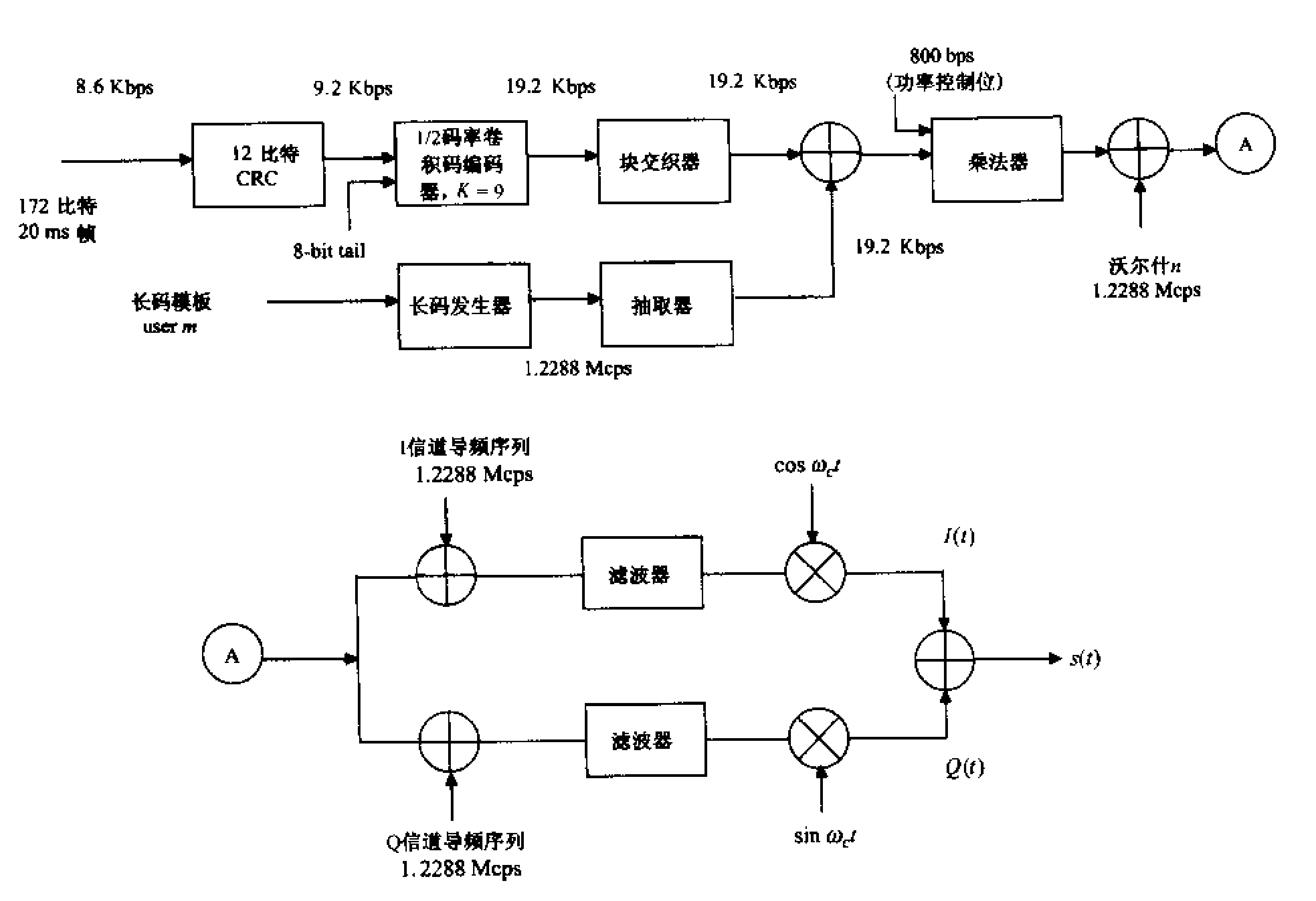


图8.19　IS-95前向业务信道图，用户m的，全速率情况

下面详细讨论这两个图,依次集中在每一个块上。首先,注意到图 8.18 和图 8.19 所示的都是全速率通信传输,速率为 8.6 Kbps,显示在每个图的左上角。IS-95 中同时定义了降低了的传输速率 4.0 Kbps,2.0 Kbps 和 0.8 Kbps,同样显示在每个图中的左上角。对于全速率情况,IS-95 定义了连续 172 比特通信帧,20 ms 长,这相当于信息输入速率为 8.6 Kbps。每帧附加了 12 个前向差错控制比特。在图 8.18 的反向业务信道情况下,输出的 9.2 Kbps 信道送给码率为 1/3,约束长度 $K=9$ 的卷积码编码器,每帧附加了 8 个全 0 尾比特[注意到对于图 8.19 中的前向业务信道情况,卷积码的码率为 1/2。此编码器提供了一些与码率为 1/3 时相比较小的差错控制能力(Viterbi,1995)[①]。为什么可以用于前向或者下行方向?]。28.8 Kbps 卷积码编码器输出发送给块交织器,减小了持续很多比特的突发错误的影响,突发错误由于移动台经历严重的衰落而产生。块交织器以每帧 576 比特(28.8 Kbps 乘以 20 ms)连续工作。块交织器可看做是一个 32 行 18 列的矩阵,如图 8.20 所示。交织器填满了每一帧,一次一列,然后逐比特读出,一次一行:依次为比特 1,33,65,97,…,545;2,34,66,…,546;…。这个过程减小了突发错误可能带来的不良影响。

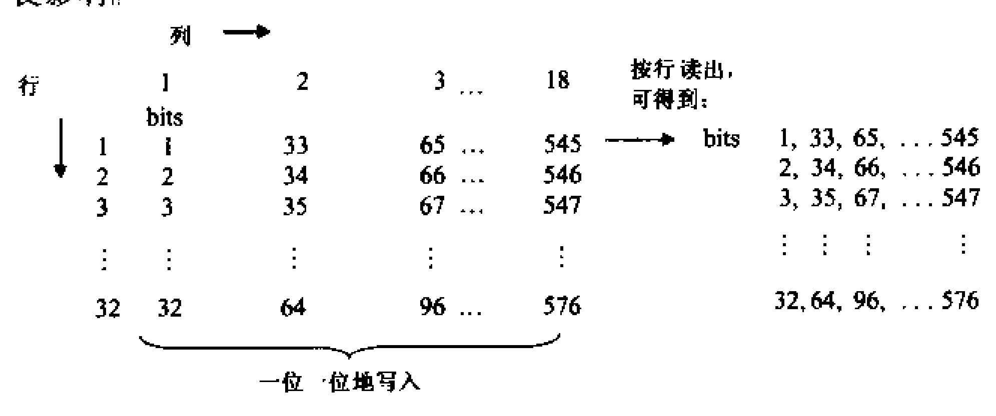


图 8.20　块交织器的交织过程

块交织器输出速率为 28.8 Kbps,发送给 64 Walsh 编码器,如图 8.18 所示。该装置需要侧面讨论。28.8 Kbps 数据流中的 6 个连续比特被存储,用于产生 64 个正交 Walsh 函数,一组二进制波形序列,每一个 64 比特长。28.8 Kbps 输入数据流被转换为 307.2 Kbps 输出数据流。为什么使用这个过程? Walsh 编码器作为一个正交调制器,与使用 64 个正交的正弦函数类似。但是使用二进制波形,这个过程接收端可以在解交织和译码以后使用非相干解调,恢复原始发送信息。直接二进制 PSK 提供更好的性能,但是需要相干解调和相位参考,例如需要和通常信号一起传输导频信号。考虑到 IS-95 中使用的无线信道相干等级,这显然是不切实际的,因为需要相对较大的能量发射导频信号。差分 PSK,例如 IS-136 系统中使用的,提供比本系统使用的 Walsh 正交策略较差的性能。

64 Walsh 编码器产生的过程由下面的矩阵描述。首先定义 Walsh 矩阵 $\boldsymbol{W}_2$ 为 2×2 矩阵

$$\boldsymbol{W}_2=\begin{bmatrix}0 & 0\\0 & 1\end{bmatrix}\tag{8.1}$$

下面用 $L/2\times L/2$ 的 Walsh 矩阵 $\boldsymbol{W}_{L/2}$ 定义 $L\times L$ 的 Walsh 矩阵 $\boldsymbol{W}_L$ 如下:

① Viterbi, A. J. 1995. *CDMA, Principles of Spread Spectrum Communication*, Reading, MA, Addison-Wesley.

$$W_L = \begin{vmatrix} W_{L/2} & W_{L/2} \\ W_{L/2} & W'_{L/2} \end{vmatrix} \tag{8.2}$$

在这里矩阵 $W'_{L/2}$ 中相应的元素 w'_{ij} 与矩阵 $W_{L/2}$ 中的元素 w_{ij} 互补,例如 $w'_{ij} = 1 - w_{ij}$。特别地,Walsh 码 W_4 为

$$W_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \end{bmatrix} \tag{8.3}$$

[同样,可以等效为最初定义矩阵 W_1 只有一个元素 0,然后反复递归地使用式(8.2),可以依次得到式(8.1)和式(8.2)]

Walsh 码 W_8 为

$$W_8 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} \tag{8.4}$$

注意到每一行有 $L/2$ 个(这里是 4 个)位置不同。同时,注意到每一行有一半($L/2$)0 和一半 1。如果将每一个 0 转换为 -1,而 1 保持不变,则从式(8.4)中可以看出,将不同两行的同一列元素相乘,再按列加起来,得到的结果为 0。特别地,

$$\sum_k w_{ik} w_{jk} = 0 \tag{8.5}$$

在这里 w_{ik} 为矩阵第 i 行第 k 列的元素。求和遍历所有列,所有列相互正交。持续这个迭代过程,用 W_8 表示 W_{16},用 W_{16} 表示 W_{32},最后用 W_{32} 表示 W_{64},所有这些矩阵具有相同的属性,如前所述,一半元素为 0,另一半为 1,所有行有 $L/2$ 个位置不同,全部满足式(8.5)中的正交关系。细节留给读者思考。特别地,对于 W_{64} 矩阵,64 个不同的行,每一个 64 比特用于提供 64 Walsh 码,用于实现图 8.18 中的反向业务信道的 Walsh 编码。每一个可能的 6 比特序列存储在编码器中,引起输出相应的 64 比特长的码,由合适的 Walsh 矩阵行提供。64 个码字组成一个正交集合,如式(8.5)所示。

下面继续讨论图 8.18 中的反向业务信道,注意到通信信号所需的 CDMA 扩频可通过对 1.2288 Mcps 的伪噪声码序列和 307.2 Kbps Walsh 码序列进行模 2 和实现。PN 码序列由 42 比特宽的长码发生器产生,如图 8.18 所示。该长码发生器包含 42 比特宽的状态发生器,具有 42 个抽头,每一个抽头的输出与 42 比特长的掩码进行模 2 组合。对产生的 42 比特进行模 2 求和,生成图 8.18 中所示的 PN 序列。该发生器与第 6 章中讨论的 42 比特长的移位寄存器相对应。长码发生器的块图如图 8.21 所示。长掩码对于系统中的每一个移动台来说是惟一的,因此相应的扩频序列可以使基站辨认移动台。掩码由 32 比特电子序列号(ESN)组成,该序列号永远分配给每一个移动终端,再加上附加的 10 比特已知序列。基站必须知道 ESN,因为在执行对移动台通过上行链路发送的 CDMA 通信信号进行译码接收管理中,它必须知道 PN 码序

列(参见图 6.5)。我们将会看到,移动台 ESN 通过上行链路传输,从移动台到基站,所有接入的信道控制信息在初始接入过程中作为移动台 ID(MSID)。

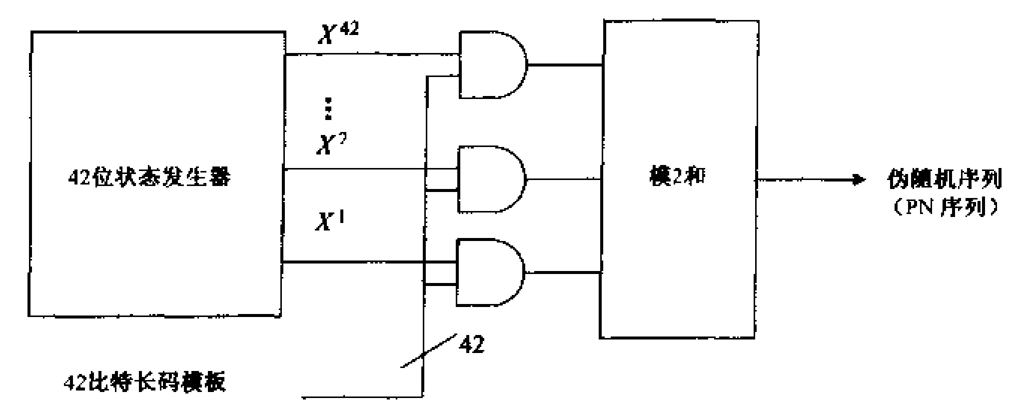


图 8.21　长码生成器

长码发生器速率为 1.2288 Mcps 的 PN 序列提供了 128 位信号扩频,这已经在第 6 章中讨论过。IS-95 系统的基带频率响应,下面将会提到,导致名义上的传输带宽为 1.23 MHz,例如,系统提供每赫兹 1 bps 的带宽。该带宽在图 8.18 中卷积码编码器的输入端被 9.6 Kbps 比特速率划分,提供扩频增益为 128。注意到长码发生器产生的 PN 序列每 $2^{42}-1$ 个码片重复一次。在 1.2288 Mcps 码片速率下,使得 PN 周期为 1000 小时,即 41.6 天! 另外,扩频同时在同相 I 信道和正交 Q 信道实现,如图 8.18 所示。这里所用的 PN 序列周期为 2^{15} 码片。在速率为 1.2288 Mcp 下,该周期相当于每 2 s 正好重复 75 次。

最后讨论反向业务信道系统框图,如图 8.18 所示。注意到 I 路和 Q 路基带低通滤波器被定义为,在 590 kHz 以内,频段具有平坦带通响应,然后下降至少 40 dB,直到截至频率起始于 740 kHz。倍增该数值得到以载波频率为中心的传输响应,这说明名义上的 1.23 MHz 传输带宽在 1.180 MHz 和 1.480 MHz 之间变化。图 8.18 所示的 D 延时窗正好为图 5.15 中的 $T/2$, OQPSK 传输所需的数值。这里的延时是码片速率 1.2288 Mcps 的一半,即 409.6 ns。

下面考虑前向(基站到移动台)业务信道的系统块图表,如图 8.19 所示。注意到该图标与刚刚讨论的图 8.18 中的反向业务信道类似,但有一些明显不同的地方。注意到卷积码编码器的码率为 1/2,而非反向信道中的 1/3 码率。对用户定义的 ESN 在 19.2 Kbps 速率下进行模 2 和运算,以提供数据的不规则性:来自长码生成器的每 64 个 PN 码片中的第 1 个与块交织器的输出进行模 2 和运算。但是,宽带扩频信号通过与 Walsh 码的码片在 1.2288 Mcps 速率下进行模 2 和运算得到。这里 64 Walsh 码用法与反向链路的情况一样,但是 64 种可能的 Walsh 码数值用于不同的目的:它们用于指示传输信息的业务信道或者控制信道。在图 8.19 的情况下,例如,基站通过业务信道 n 发送给用户 m 个信息,指定为 Walsh 码 W_n。这种利用码字指示信道的技术当然与前面讨论的用于 TDMA 系统、GSM 系统和 IS-136 系统中的时隙信道技术不同。

64 种可能的 Walsh 码,用于定义前向通信或控制信道,具体描述如下:Walsh 码 W_0,包含全 0 序列,与导频信道相对应,为移动台提供所需的载波相位和在给定蜂窝中启动时的定时信息。同步信道用于为移动台提供附加的系统定时和系统配置信息。它用 W_{32} 表示,前面 32 比特为全 0,剩余 32 比特为全 1。每一个前向 CDMA 信道最多 7 个寻呼信道,用于接入过程的一部分,以及传输寻呼信息给移动台,由 W_1 到 W_7 表示。剩余的 Walsh 定义的信道被指定为前

向业务信道。例如,可能有 3 个寻呼信道 W_1 到 W_3,没有同步信道。在这种情况下,Walsh 信道 W_4 到 W_{63} 被指定为业务信道。然而,如果存在同步信道,W_{32} 和其他信道被指定为业务信道。如图 8.19 所示,给特定的移动台传输数据由该用户的长掩码指定,例如,长掩码序列载有该用户的 ESN。如前所示,用户 m 被这样识别,如图 8.19 所示。在这种情况下,如前所述,正在使用的特定业务信道通过 Walsh 码 W_n 进行识别。移动台使用接收频率将执行图 8.19 中所示的相反过程。细节将在后面的 IS-95 系统的控制信道中进行讨论。

注意到图 8.19 显示了一个 800 bps 的数据流,标记为电源控制比特,与块交织器中 19.2 Kbps 通信数据流复用。回忆第 6 章,CDMA 系统依赖移动台严格的功率控制,从而保证基站从蜂窝中的所有移动台接收到的功率是相同的。这通过使用每一个移动台所提供的开环控制组合来实现,被闭环校正增强,包括基站。正是 800 bps 功率控制子信道提供了该功率控制校正。基站通过 1.25 ms 时间间隔来估计每一个移动台的接收功率,然后传输一个 0 给移动台。如果该移动台的功率太低,就使其增加功率;如果该功率应该被降低,则传输一个 1。在 19.2 Kbps 信息比特速率下,24 比特每 1.25 ms 被传输一次,每一个比特持续 52.083 μs。这些比特中有两位(104.166 μs 长)每 1.25 ms 被功率控制比特替代一次,形成 800 bps 的功率控制子信道。这是一个凿孔码的例子(图 8.19 中没有画出下抽取因子为 24 的下抽取器,但实际上这个抽取器是存在的,将序列的速率从 19.2 Kbps 降到 800 Kbps,为业务信道/功率控制比特复用提供了定时控制)。

最后,注意到 I 路和 Q 路扩频序列包含导频 PN 序列标记。每一个基站根据导频 PN 序列使用一个特定的时间偏移。所有前向控制信道和给定基站的所有 CDMA 频率分配使用相同的偏移。该时间偏移指定一个给定的基站,然而偏移可以重复使用。每个时间偏移用 64 个码片指定。例如,相对于 0 偏移导频 PN 序列,偏移 10 对应 $10 \times 64 = 640$ 个码片。在码片速率为 1.2288 Mcps 的情况下,这对应于 520.83 μs 的实际延时。

下面考虑 IS-95 中使用的控制信道。已经注意到前向、下行方向具有导频、同步和寻呼信道。在这个方向上的业务信道载有功率控制子信道,以及下行链路信号信息。反向、上行方向只具有接入控制信道。所以,这与 IS-136(表 8.2)和 GSM(表 8.1)中定义的随机接入信道 RACH 类似。上行即反向业务信道同样载有信号信息。图 8.22 描述了这些不同的信道。不同的码字用于定义所有这些信道,正如前面所述的前向链路。

首先,集中讨论上行接入控制信道。如前所述,该信道在 IS-95 中的作用与 RACH 信道在 GSM 系统和 IS-136 系统中的作用相同。它是一个多址接入信道,移动台利用其开始与基站进行通信,以及回复前向寻呼链路承载的寻呼信息,如图 8.22 所示。IS-95 中定义了 32 个这样的接入信道,每一个都与特定的寻呼信道关联在一起。对于每一个寻呼信道至少有一个接入信道可用。接入信道利用自己惟一的码来辨认,如图 8.23 所示。我们只描述了与反向业务信道相应信道不同部分的图表。注意到每 20 ms 帧产生 88 个接入信道比特。8 个尾比特附加到每一帧,产生 4.8 Kbps 信道,进入码率为 1/3,约束长度 $K = 9$ 的卷积码编码器。产生的 14.4 Kbps 编码的输出逐比特重复,得到 28.8 Kbps 比特流,与图 8.18 所示的全速率业务信道的速率一样(注意到符号重复同样用于非全速率业务信道的情况,卷积码编码器输入端的速率为 4 Kbps,2 Kbps,0.8 Kbps,如图 8.18 所示。在所有情况下,28.8 Kbps 比特流被发送到这里所描述的 Walsh 编码器中,与控制信道的情况相同)。同样地,64 Walsh 正交调制在 PN 扩频以后用于反向业务信道。然而,这里不同的是 1.2288 Mcps PN 序列由 42 比特接入信道长掩码产生,长掩码由 16 比特基站 ID 组成;9 比特导频 PN 偏移值,前面提到的与基站导频 PN 序列偏

移对应的导频 PN;3 比特寻呼信道号(回顾一下,每个前向 CDMA 通道最多只能有 7 个寻呼信道可用)和 5 比特接入信道号。这 33 比特被规定的 9 比特字段扩大,在这种情况下形成 42 比特掩码。这些不同的数字和 ID,在通电以后呼叫开始时,通过前向同步信道信息和寻呼信道,所承载的系统参数信息被提供给移动台。在后面的呼叫过程中,这些信息会在所使用的系统控制信道的讨论中进行描述。

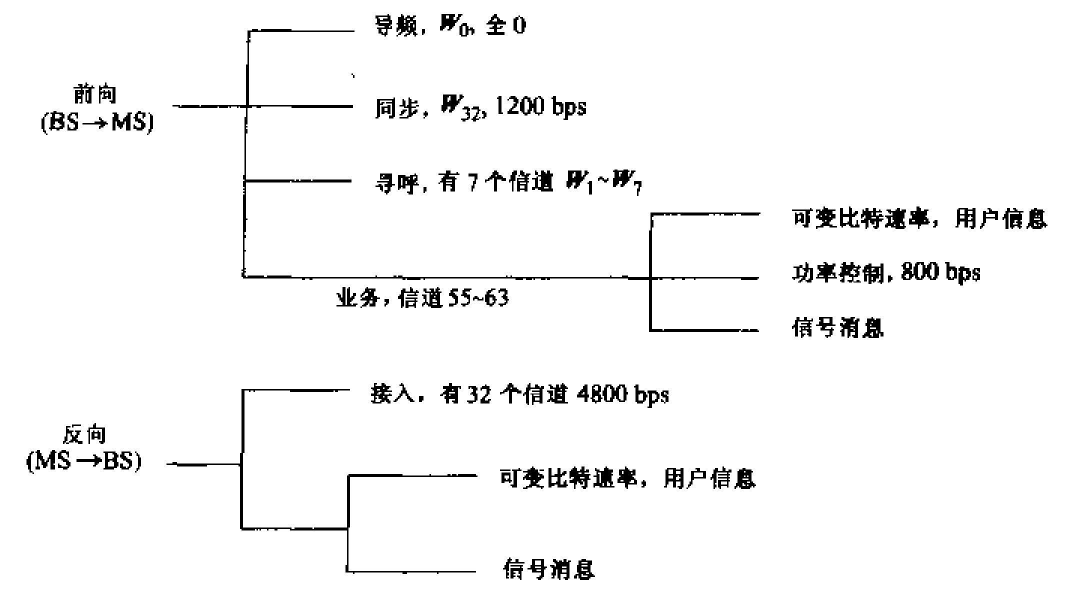


图 8.22 IS-95 的业务和控制信道

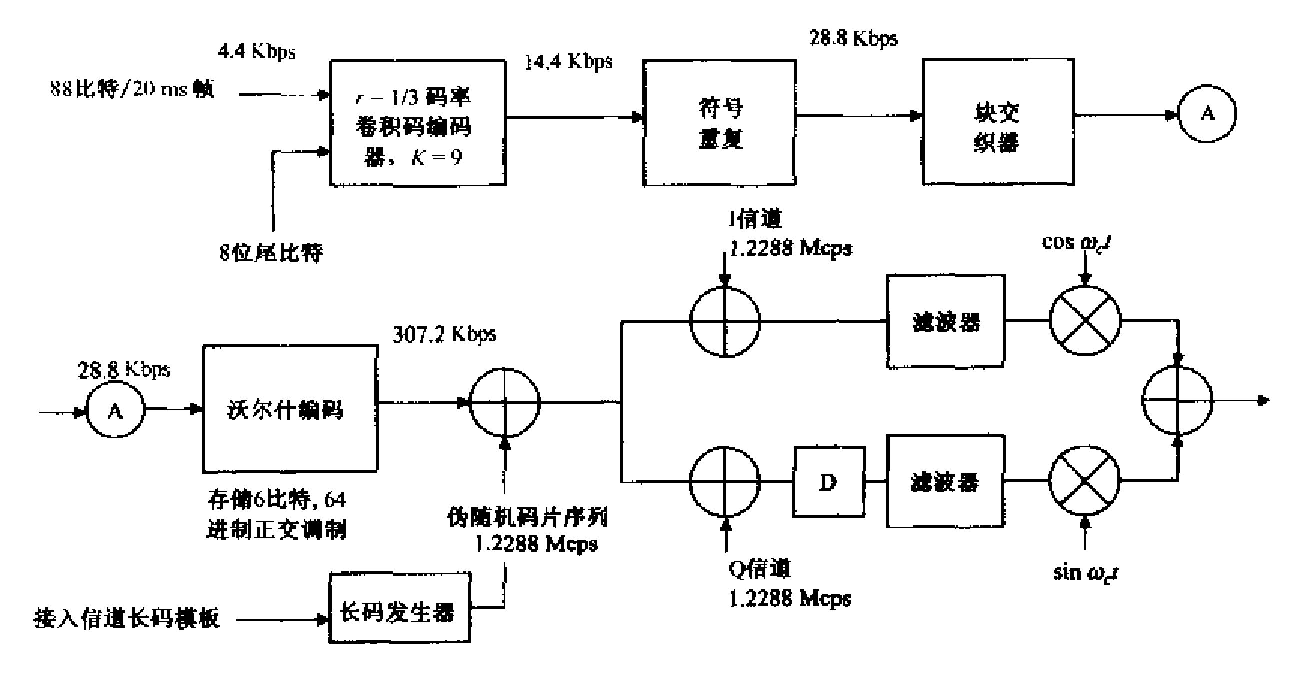


图 8.23 IS-95 的接入信道框图

修改过的 Aloha 随机接入过程用于通过接入信道给基站传输信息。正如前面讨论的 GSM 和 IS-136 系统的情况,发送该信息以回复寻呼信道接收的寻呼信息。该信息可能来自移动台,注册一个新位置,或者作为呼叫开始过程的一部分。回复尝试和请求尝试过程有一些不同,但都是利用接入信道在每一次可接入信道列表随机选择的时间上进行一系列接入尝试。每一次接入尝试实际上由很多接入探测器组成。每一个探测器增加功率,就和前一个探测器随机延

迟一段时间。随机退避延时也被用于每一次多探测器接入尝试之间。在接收到寻呼信道的确认时,接入可以在序列的任何时间完成。接入信道信息都载有 34 比特的移动台信息号(MIN)和 32 位 ESN,这在前面已经提到过。

图 8.22 所示的前向(基站到移动台)控制信道描绘在图 8.24 中。如图所示,全 0 导频信道被全 0 Walsh 码 W_0 扩频,扩频速率为 1.2288 Mcps。同步信道工作在 1.2 Kbps 速率,通过码率为 1/2 的卷积编码和符号重复变为 4.8 Kbps。块交织以后,同步信道比特通过添加与鉴别信道,对速率为 1.2288 Mcps 的 Walsh 码 W_{32} 模 2 和结果进行扩频。寻呼信道,最多有 7 个,初始速率为 9.6 Kbps 或 4.8 Kbps 的输入信息,通过卷积编码和对 4.8 Kbps 信号进行符号重复变为 19.2 Kbps,根据分配的寻呼信道的数量,通过 1.2288 McpsWalsh 码,从 W_1 到 W_7 进行扩频。在图 8.19 中,所有这些控制信道和前向业务信道,Walsh 码都是 64 码片长,每 64 个码片重复一次,码片速率为 1.2288 Mcps。注意到所有这些控制信道的 I 路和 Q 路扩频信号具有相同的基站导频 PN 序列偏移。

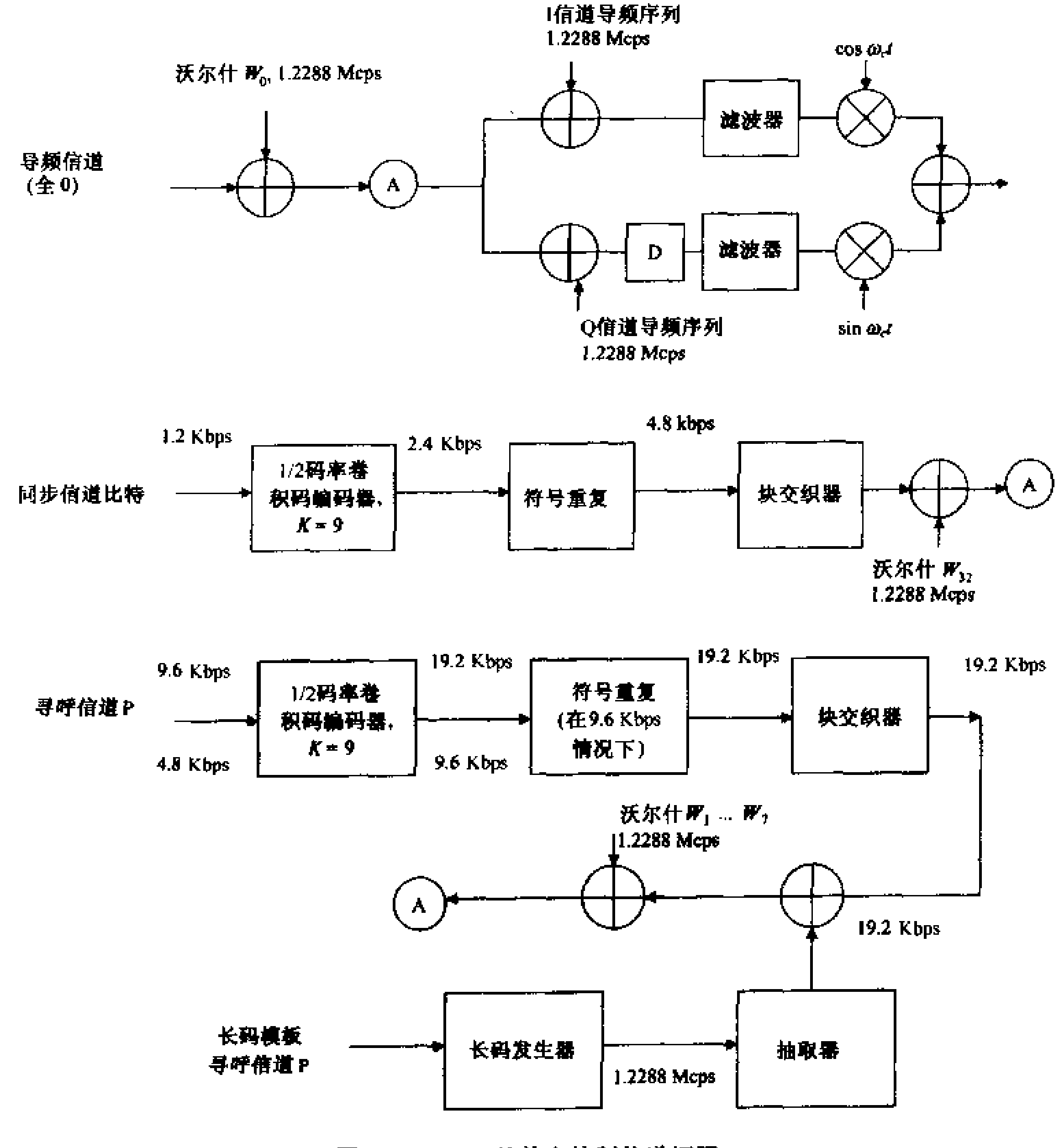


图 8.24 IS-95 的前向控制信道框图

如何利用这些不同的前向控制信道呢？导频信道用于给定蜂窝中的移动终端，前面已经提到，获取载波相位和定时参考。同步信道工作在 1200 bps，只传输一个信息，包含系统时间、9 比特偏移值、系统/网络 ID 和所用的寻呼信道信息速率 4800 bps 或 9600 bps。寻呼信道用于给没有分配到业务信道的移动台传输控制信息。一个明显的寻呼信道信息的例子是特定终端的寻呼信息，可通知另一个将要开始一次呼叫的用户。这个过程将在下一章详细讨论。寻呼信道也用做接入控制协议的一部分，确认接收到移动台产生的接入信道信息。寻呼信道发送的其他信息包含信道分配信息，在讨论呼叫连接过程中会简要提到它；系统参数信息，载有基站 ID、系统/网络 ID、Pilot_PN 和该 CDMA 信道中可用于寻呼信道的数量以及其他信息；接入参数信息，包含移动台的功率信息、可接入信道的接入信道数量、Pilot _PN，以及接入过程有关的信息，例如，接入信息格式以及其他一些参数；邻居表信息；CDMA 信道列表信息；可变顺序信息和很多其他类型的信息。在这些信息中有一些被周期性地发送，其他根据需要决定是否被周期性地发送。我们将会在讨论呼叫开始时和举例说明这些信息实际上是如何通过寻呼信道被传输时进一步进行讨论。它们将被分为提供系统和接入参数的广播信息、呼叫管理信息、认证信息，以及管理和维护信息几类(Goodman; 1997)。

IS-95 中用于建立呼叫的过程同时使用了前面讨论的反向和前向控制信道，这和前面图 8.3 所描绘的 GSM 系统与图 8.15 所描述的 IS-136 系统类似。在给定蜂窝中，打开电源的移动台首先按照图 8.25 所示的过程工作。为全 0 导频信道设置接收码信道，为该蜂窝寻找和获取导频信道，然后为同步信道设置接收码信道并获取该信道。此时移动台处于空闲状态，监测寻呼信道，以便获取开始呼叫所需的系统和接入信道参数信息。这些参数包括正在使用的特定寻呼信道，以及与该寻呼信道关联的相应的接入信道。一旦获取到这些参数，移动台就准备利用接入信道开始呼叫过程。

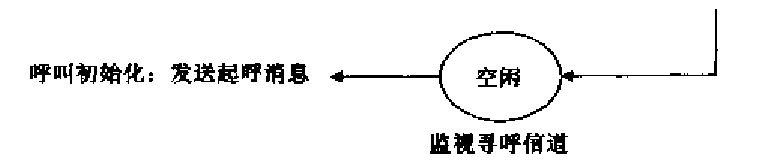


图 8.25　IS-95 中建立呼叫的准备过程

从空闲状态，移动台也可以通过发送注册信息利用接入信道在基站注册。该信息包含移动台 ID 和其他字段，为系统指示移动台位置。注册以后，移动台返回空闲状态或者开始呼叫过程，也可以回复等待的寻呼信息。

下面讨论呼叫开始的过程，该过程指示传输信息和承载它们的信道。图 8.26 描述了这个过程，与图 8.3 和图 8.15 比较，分别显示出与 GSM 和 IS-136 系统类似的呼叫建立过程。这里使用的信道是码字，而非基于 TDMA 的在 GSM 和 IS-136 系统中使用的时隙。前向、下行方向由 64 种可能的 Walsh 中适当的一种指定(参见图 8.19、图 8.22 和图 8.24)，它们的长码掩码作为反向、下行方向 PN 码片序列的一部分(参见图 8.18，图 8.23)。如图 8.26 所示，移动台通过向接入信道发送开始信息开始呼叫连接。假设这个信息被正确接收，基站通过向寻呼信道发送信道分配信息进行回复。该信息为移动台指示基站进行传输即将使用的前向业务信道，其业务信道接收码被设置。接收到信道分配信息的移动台开始通过反向业务信道发送导言，包

含速率为 9600 bps 的连续 0,用来使基站确认该信道。此时,基站通过反向通信信道发送基站确认序列,然后业务信道中的双向通信就可以开始了。

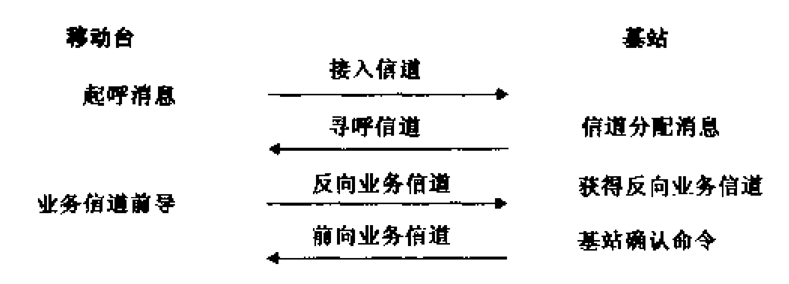


图 8.26 IS-95 简化的呼叫发起过程

这些信号和控制信息的格式是什么?考虑方向情况。注意到它们被装载在接入信道或反向业务信道中。首先,考虑发送到反向业务信道(数据通信载波)中的信号信息,这是第二代蜂窝系统中主要的语音信息。已经指出,反向业务信道可以在 9600 bps、4800 bps 和 2400 bps 的速率下工作。如果需要的话,可以逐帧变化。为了简化,只集中讨论全速率 9600 bps 的情况,如图 8.18 所示的反向业务信道系统的框图也是如此。连续反向业务信道帧 172 比特长,每 20 ms 重复一次,可能载有业务数据、主要标志业务和信号业务的混合,也可能载有可选的次要业务。表 8.3 指示图 8.18 中每帧 172 比特是如何在数据和信号或次要业务之间进行分配的。如表 8.3 所示,帧中第 1 个比特(混合模式比特),如果被置为 0,表示给定帧仅载有主要业务;如果被置为 1,表示该帧载有主要业务和信号或者可选次要业务的混合。接下来的 3 个比特,如果混合比特被置为 1,表示剩下的 168 比特在主要业务和信号或次要业务之间的划分。载有 16 比特主要业务和 152 比特信号业务的帧的例子如图 8.27 所示。术语 dim-and-burst 用于指示那些同时载有主要业务和信号业务的帧。Blank-and-burst 专门用来指示载有信号业务的帧。8.5 节将讨论如何同时使用反向和前向业务信道处理语音呼叫,所以这里仅集中讨论信号消息。

表 8.3 IS-95 反向业务信道,172 比特帧(9600 bps 的情况)

比特每帧: 格式域比特 混合模式	业务类型	业务模式	主要业务	信号业务	次要业务(可选)
0	–	–	171	0	0
1	0	00	80	88 *	0
1	0	01	40	128 *	0
1	0	10	16	152 *	0
1	0	11	0	168 * *	0
1	1	00	80	0	88
1	1	01	40	0	128
1	1	10	16	0	152
1	1	11	0	0	168

注:* 主要业务和信号业务混合
* * 只有信号业务

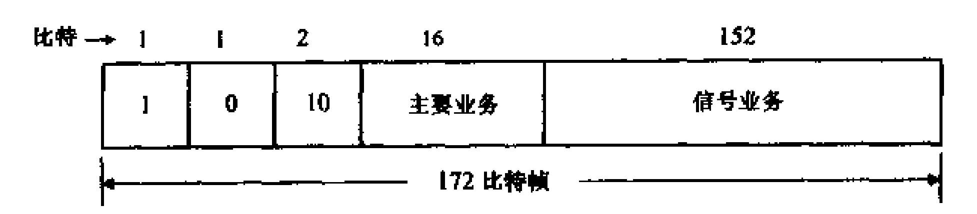


图 8.27 IS-95 中反向业务信道"主要业务和信号业务混合"帧的例子

发送到反向业务信道上的信号消息长度在 16 比特到 2016 比特范围内变化，所以，它们可以根据消息长度以及数据和信号业务之间的帧的划分情况占用复帧。然而，在被分成复帧之间，消息具有 8 比特长度的首部，同时附加 16 比特的差错校验序列。格式如图 8.28 所示。消息长度指示器指示消息长度，单位为字节。图 8.29 给出了一个例子。在这个例子中，信号消息主体被任意选择为 37 比特到 296 比特长。全部消息被装载在帧中，包含 24 比特首部和 CRC，总共 320 比特长。每帧具有 80 比特的主要业务和 88 比特的信号业务的帧的 dim-and-burst 划分被任意选取，所以需要用 4 个帧来际载这个信息。在每帧 88 个信号比特中的第 1 个比特用于指示当前帧是否为信息或持续帧的第 1 帧。每帧剩余的 87 比特载有信号消息。261 个比特被装载在前面 3 帧；剩余 59 比特被装载在第 4 帧和最后 1 帧。该帧中剩下的 28 比特必须被全 0 填充比特填满。基站接收到该消息以后，使用消息长度指示器删除填充比特。



图 8.28　IS-95 反向业务信道信号消息格式

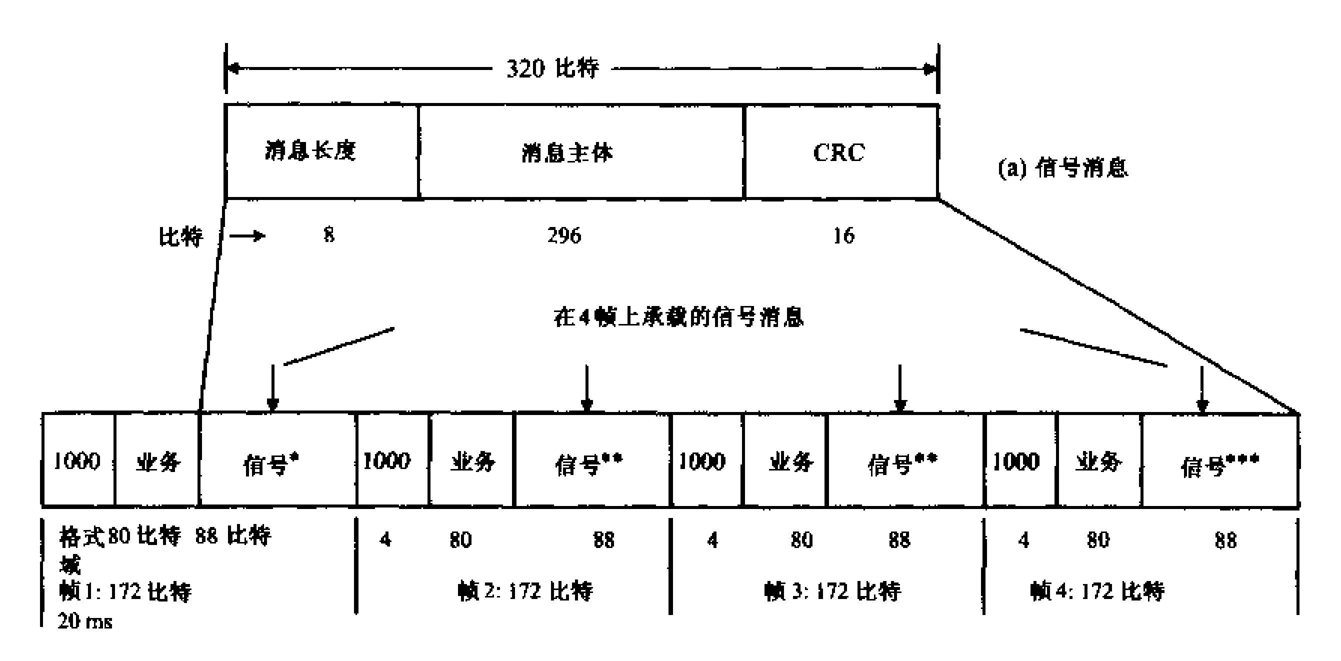


图 8.29　IS-95 中反向业务信道，多帧信号消息的例子

IS-95 中定义了很多反向业务信号信息，每一个根据发送的消息都是长度可变的。例子包括导频强度测量消息，它为基站提供了 6 比特的移动台导频信号功率测量的指示，该导频信号功率可用于计算移动台的时间参考。功率测量报告信息给基站提供了移动台为每一个导频测量的功率，以及用于进行这个测量的帧的数量；开始连续消息包含任何没有包括在接入信道发送开始消息之内的拨号数字；信号消息标志包含信息记录，例如，呼叫组和被呼叫组的号码；状

态消息包含记录信息，例如，移动台标识、ESN 和站点 MIN、包含制造商代码和模型号的终端信息，以及移动台的安全状况；一些命令消息、典型的简单消息，例如释放和连接命令。当发送的时候，这些都将按照已经描述过的格式进行发送，表 8.3、图 8.28 和图 8.29 中的信号业务部分对此进行了总结。

前面的讨论集中在反向业务信道载有的信号消息。现在考虑接入信道载有的信号消息。如前所述，该反向控制信道用于移动台建立呼叫(参见图 8.26)、回复寻呼或者注册它的位置。3 种不同定义的接入信道消息，包含许多命令消息。分别实现这些不同工作，是前面已经提到的起呼消息、寻呼响应消息和注册消息。所有这些消息有移动台 ID 和其他信息。消息长度在 2 比特到 842 比特之间变化。它们被嵌入到接入信道消息群体，长度为整数个 88 比特接入帧的长度(回想图 8.23，20 ms 的接入帧载有 88 个比特)。图 8.30 显示了一个接入信道消息。注意到消息主体前面是 8 比特用于指示长度的消息长度字段，包含 30 比特差错校验字段(但不包括所显示的填充)跟随在消息主体之后。群体可以不多于 800 比特或 10 帧。群体最小长度是 3 帧。因为群体必须与帧的边界一致，必须用填充增加帧的长度，如图 8.30 所示。最大群体长度由基站决定，当移动台处于空闲状态时，通过寻呼信道使用接入参数消息发给移动台(参见图 8.25)。群体中每一个 88 比特帧通过 8 个全 0 尾比特增长，在进行卷积编码前形成 96 比特帧，如图 8.23 所示。

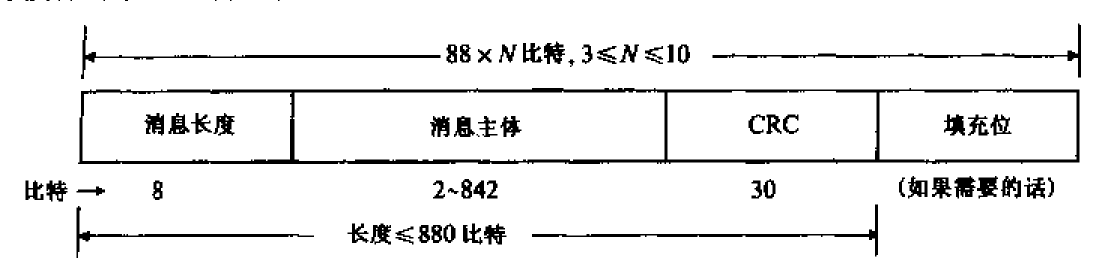


图 8.30 IS-95 接入信道消息结构

接入信道传输，称为接入信道时隙，包括增长的具有全 0 比特前导的消息群体，最小长度为 1 帧或 96 比特。基站用前导获取或通过接入相位与移动台传输同步。前导长度最大为 16 个全 0 的 96 比特帧，也由基站在寻呼信道发送的接入参数中指定。图 8.31 描述了一系列接入信道时隙，每一个包含 3 帧前导和 4 帧群体。传输以 4800 bps 速率输入卷积码编码器，如图 8.23 所示。

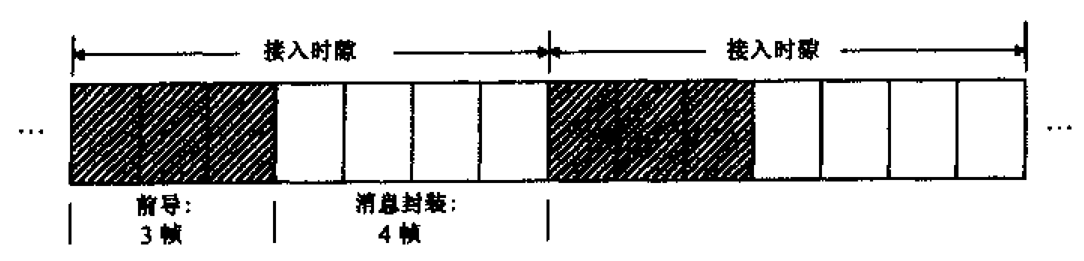


图 8.31 IS-95 接入信道，传输时隙，4800 bps

到目前为止，我们集中讨论了在反向信道(从基站到移动台)上发送的信号和控制消息。前向、下行方向发送的信号和控制消息的格式是什么？回忆图 8.22，下行方向的信号消息可能通过前向业务信道、3 个前向信号信道，以及前面提到过的同步、导频和寻呼信道进行发送。前向业务信道中的信号与前面讨论的反向业务信道中的信号很类似：前向业务信道由 172 比特帧组织。这些帧载有主要业务消息、信号业务消息和次要业务消息的混合，与反向通信帧组

织形式一样，如表 8.3 所示，例子如图 8.27 所示。对于前向业务信道（从基站到移动台）信号消息格式本质上与图 8.28 所示的反向业务信道情况相同，只有一个差别，消息主体可以不超过 1160 比特。下行链路信号消息需要复帧，通过图 8.29 所示的反向业务信道类似的前向业务信道发送。IS-95 中定义了很多前向业务信道的消息。例子包括很多命令消息、认证挑衅回复消息、切换方向消息、邻居表更新消息和功率控制参数消息，以及其他消息。

现在考虑前向信号信道。如前所述，导频信道用于发送提供给移动台所需的载波相位和定时参考的全 0 序列（Walsh 码 W_0）。同步信道消息，如果由基站来发送，则是 162 比特定长消息，通过同步信道利用 Walsh 码 W_{32} 持续发送，这在前面已经提到过（参见图 8.24）。同样，如前所述，该消息包含 15 比特系统 ID、16 比特网络 ID、每个基站惟一的 9 比特导频 PN 序列偏移、36 比特系统时间参数和寻呼信道比特速率指示器，以及其他参数。传输该消息所用的格式与业务信道消息所用的格式类似（参见图 8.28）：162 比特消息被装载在 279 比特群体中，另外包含 8 比特消息长度字段、30 比特 CRC 字段和 79 比特填充，一共 279 比特，如图 8.32(a) 所示。群体被分成 93 比特片断，每一个片断作为由 3 个 32 比特连续帧构成的复帧传输，每一帧包含 31 比特，前面有 1 比特的架构比特：“1”用于复帧组的第 1 帧，“0”用于其他两帧[参见图 8.32(b)]。每 80 ms 传输一个这样的复帧，如图 8.32 所示。注意到这将导致图 8.24 所示的同步信道输入传输速率为 1200 bps。3 个复帧用于传输每一个同步信道消息，如图 8.32(b) 所示。

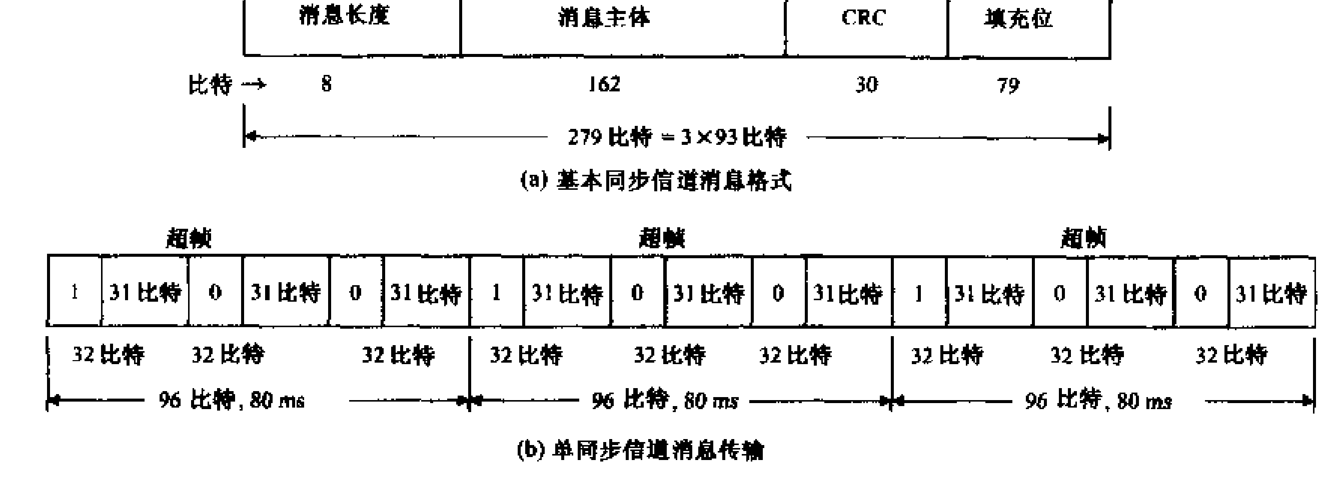


图 8.32 IS-95 同步信道消息传输

寻呼信道用于传输在很多业务信道分配前从基站到移动台的控制消息，如前所述。我们也要涉及到图 8.26 所示的用于呼叫开始过程的信道分配消息、系统参数消息和接入参数消息等其他寻呼消息。寻呼信道消息发送速率为 4800 bps 或 9600 bps（参见图 8.24）。传输被组织成 80 ms 时隙，每一个时隙根据使用的传输速率载有 384 比特或 768 比特的数据。时隙被组织成 2048 个163.84 s 长的时隙循环，允许移动终端定期进入休眠模式或空闲状态，然后在一个循环的特定时间内被周期性唤醒。这个过程用于保存电池能量。每一个 80 ms 时隙被分成 4 个 20 ms 帧，每一帧包含两个 10 ms 半帧。半帧长度为 48 比特或 96 比特，第 1 个比特为 SCI (synchronized capsule indicator) 比特。该比特指示消息传输是否同步。这些半帧载有的寻呼信道消息，每一个都包含一个消息主体，前面有 8 比特长的消息指示器，后面有 30 比特长的 CRC。消息主体，如图 8.33 所示，长度在 2 比特到 1146 比特范围内变化。可能附加填充比特

形成消息群体。如果使用同步传输,则同步消息群体从半帧的第 2 个比特(紧随 SCI 比特)开始,用填充进行保证。

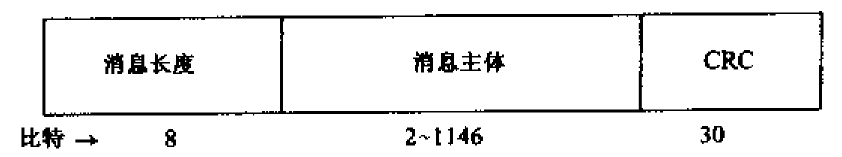


图 8.33 寻呼信道消息格式

8.4 移动管理:切换、定位和寻呼过程

在这章的引言中,涉及到当移动台在无线网络中移动或者漫游到其他系统时,通过移动台和基站之间的空中接口发送信号,以及通过有线网络发送信号来控制移动台通信。到目前为止,本章集中讨论了通过空中接口发送信号,给出当前第二代系统 GSM,IS-136 即 D-AMPS 和 IS-95 使用的 3 种主要标准的例子。本章前面 3 节已经详细描述了建立呼叫和实现其他必需的控制功能所需的不同的信号消息,以及如何通过移动台和基站之间的空中接口进行传输。本节集中讨论当移动台在蜂窝之间或在多个蜂窝内大范围移动时所需的信号。处理移动台移动所需的控制称为移动管理。移动台的移动本质上包括 3 种功能:切换控制、定位管理和寻呼。切换控制是当移动台从一个蜂窝移动到一个临近蜂窝或者从一个系统移动到另一个系统时正在进行的呼叫所必需的控制。一组新的传输信道必须被分配给切换呼叫,不能让用户感觉到呼叫的切换或者中断呼叫。两种类型的切换分别为蜂窝间切换和系统间切换。当移动台/基站共同决定功率或信道衰落保证在蜂窝内变换信道时,也可能有蜂窝内切换。

定位管理用于处理移动台处于外地区域的注册,使其在呼叫到来时能够被定位和寻呼。回忆本章引言中图 8.1 所示的与本地区域关联的移动台,其所需的参考和侧面信息被存储在本地定位注册数据中,如图 8.1 所示的 HLR 区域。每一个 HLR 与一个移动台切换中心关联,如图 8.1 所示。定位或注册地点用于指定多个蜂窝的范围,每一个都带有自己的基站,由 MSC 控制。在不同区域注册,访问位置寄存器 VLR 在本地地点的 MSC 和 HLR 之间进行通信之后,用于存储来访移动台的参考和侧面信息。这个注册信息首先询问本地 MSC 和 HLR,然后用于在呼叫到来时寻呼移动台。通过访问 MSC 给其控制的所有基站发送寻呼消息,然后向它们的寻呼信道广播寻呼消息。

在 MSC 之间发送的信号消息执行不同的移动管理功能,前面已经暗示过,是通过有线网络传输。在北美,电信产业联合会在 IS-41 中定义了用于处理系统之间功能的各种协议和消息描述。在 IS-41 中假设潜在的数据网络用于承载系统之间的执行移动台管理功能的信号消息。在实际中,所使用的数据网络遵循另外一套称为信号系统数字 7,简称 SS7 的标准。SS7 是一个分层协议结构,网络层包括数据交换,在全球范围内为固定电话网络提供呼叫管理功能,所以很自然地采用了这个标准和它为移动电话管理定义的消息集。关于 SS7 的指南可参见 Modaressi 和 Skoog(1990)的著作。[1]

现在进一步详细研究移动管理 3 个部分中的每一个。首先研究切换控制,先讨论蜂窝之

---

[1] Modaressi, A. R. and R. S. Skoog. 1990. "Signaling system number 7: a tutorial," *IEEE Communications Magazine*, 28, 7(July), 19-35.

间的切换，然后讨论系统之间的切换，再然后讨论定位管理和寻呼，包括简单考虑处理这两个矛盾功能的折中办法(移动终端之间交叉范围的下界出现在定位范围尺寸增加、注册数量减少，但广播更多的寻呼消息去寻找给定的移动台的情况下)。

考虑正在通信的移动台在给定蜂窝内的移动。当移动台到达蜂窝边界时，正在与移动台通过空中接口进行通信的蜂窝基站接收到的功率将降低到预先定义的门限值以下。相比之下，假设从临近基站接收到的功率超过了门限值。MSC 通过控制这两个基站，决定切换到邻近基站，进入与该基站关联的新蜂窝。第二代系统支持移动台协助切换(MAHO)，功率测量通过移动台在 MSC 和基站的命令来实现，测量结果传输给基站作为上行信号消息。

以 IS-136 系统作为一个具体的例子。GSM 使用类似的切换策略。图 8.34 给出了当执行移动台协助切换时在 MSC、基站和移动台之间传输的消息。MSC 通知基站进行信道质量测量。基站通过给移动台传输测量命令消息进行回复，该消息识别出邻近的前向业务信道和当前从基站接收到的消息，该消息用于测量信道质量的前向业务信道。信道质量测量包含当前业务信道和邻近业务信道的接收信号强度测量，以及当前信道的比特差错率测量。这些测量结果完成以后，移动台通过将信道质量测量消息装载到 SACCH 中，把消息反馈给基站(见 8.2 节)。然后，基站把测量结果传给 MSC，发布停止测量的命令，再通过基站向移动台发送停止测量的命令消息。根据测量分析，如果切换确实需要，MSC 将命令基站，然后给移动台发送信号(图 8.34 示意了从移动台上行链路发送应答信息，应答纠正接收到对应的命令信息)。

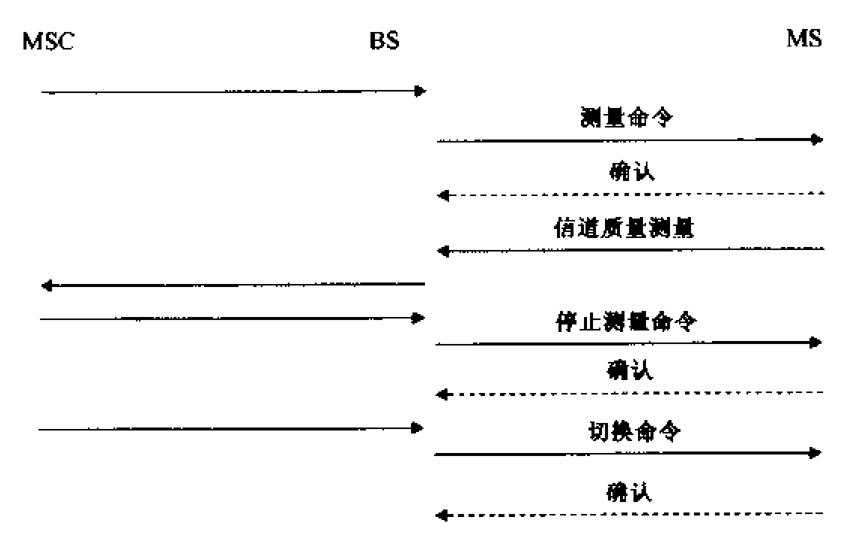


图 8.34 辅助移动切换，IS-136

移动台切换到一个新基站必须立即给移动台在新蜂窝中分配信道。如果没有可以利用的信道，则正在进行的通信将被中断。如果在分配信道时新呼叫和切换被同样对待，则切换带来的中断概率明显与阻止一个新产生的呼叫的概率相同。阻止呼叫的概率在第 3 章中讨论过[见式(3.7)]。可以肯定的是，移动系统用户发现正在进行的呼叫中断比在尝试开始呼叫时接收到忙信号的情况严重。毕竟，用户不希望在通信的过程中被中断！所以人们提出很多建议降低中断呼叫的概率。这些包括相对于新呼叫阻断而言给予切换呼叫优先权，或者当等待信道可用时用简短的时间询问切换呼叫。在信道可用数固定的情况下，给予切换呼叫优先权明显增加了阻塞性呼叫的概率，所以必须折中考虑。下一章在研究呼叫准许控制时我们再回来考虑这些建议，如果可能的话，再进行比较。

第6章提到基于CDMA的IS-95系统使用软切换而非硬切换，我们已经暗含假设了这一点。注意到那里的软切换导致了信干比增益，进而提高了系统性能。在软切换过程中，移动台在选择与哪一个基站通信前，与两个或多个基站连接，所以，这是一个“中断前先连接”的例子。硬切换暗示在连接到新基站之前，断开与当前基站的连接。IS-95系统的切换过程使用移动台协助切换，IS-136和GSM系统也是如此。但是与那些采用硬切换过程的系统不同，在IS-95中，移动台发起切换过程。移动台根据自身进行的测量决定是否需要切换。然而，由MSC进行实际的切换决定。当MSC通过给基站发送消息命令移动台中断与原来基站的连接，并维持与新基站的连接时，切换才会完成。但是，与原来基站和新基站的连接一直维持到已连接到新基站的决定做出以后，这是软切换的本质。

现在考虑系统之间的切换。这种类型的切换用于移动台漫游，当移动台从一个MSC控制的区域移动到另一个MSC控制的区域。IS-41，北美系统间移动通信标准定义了很多可能的切换情形。在图8.35中，描述了系统间切换的简化情况，移动台 m 从一个MSC控制的区域，这里标记为MSC-A，移动到MSC-B控制的区域，这个过程称为前向切换。当移动台移动穿过区域A时，如图8.35(a)所示，它最终进入A和B之间的重叠区域[参见图8.35(b)]。

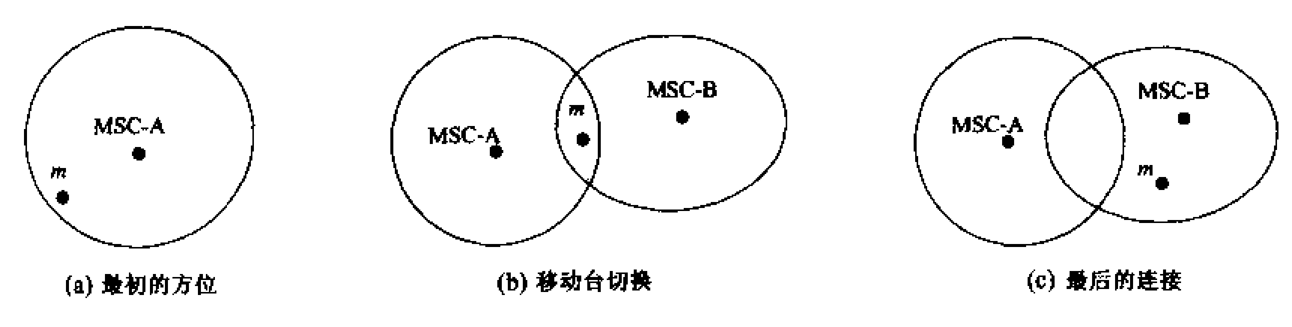


图8.35 IS-41系统间的前向切换

这里MSC-A根据它所做的测量决定把呼叫转给了MSC-B。一旦呼叫转换完成，移动台由MSC-B控制，如图8.35(c)所示。当移动台在图8.35所示的重叠区域时，这两个MSC发送和接收的消息如图8.36所示。标记调用(INVOKE)和结果返回(RETURN RESULT)代表SS7消息类型，用于承载IS-41消息。最初处理移动台 m 呼叫的MSC-A决定了切换到邻近MSC-B是否合适。然后，它发送切换测量请求INVOKE消息给MSC-B，其中载有MSC-A中的服务蜂窝的ID。MSC-B用一个或多个蜂窝的列表进行回复，其中包括信号质量信息。然后，MSC-A决定是否切换。如果决定切换，则它在两个MSC之间建立一条电路，然后给MSC-B发送指示设备调用消息(FacilitiesDirective INVOKE)，则载有MSC-B的蜂窝ID就分配给了蜂窝处理移动台。如果MSC-B在指定蜂窝内发现一个可用的语音信道，则它用结果返回消息回复，然后MSC-A命令移动台 m 转换到该语音信道。MSC-B向移动台所在的信道发送设备释放调用消息(Facilities-Release INVOKE)，释放MSC之间的电路，则切换过程完成。

IS-41中定义的其他系统之间的切换过程包括初始切换返回原来的MSC或者转向第3个MSC的情况。

现在简要描述定位管理。回忆到定位管理涉及正在漫游的移动台，在它们经过的任何新地点进行注册的请求。然后，它们可以在呼叫到来时被寻呼。包含多个蜂窝的区域由MSC来控制。讨论IS-41中描述的北美标准。GSM中指定的过程与其非常类似。与每一个MSC服务区域相联系，我们经常将定位管理认为是位置或注册区域，与其相关联的是访问位置注册

(VLR),它可维护在本地注册的异地移动台的数据。给定区域的基站周期性地广播区域 ID。对于正在漫游的移动台,我们通过监听基站广播,来探测它是否已经进入到一个新区域。一旦它进入到一个新区域,就开始注册过程。图 8.37 简要描述了这个过程。图中假设移动台原来在 MSC-1 控制的区域,与 VLR1 关联。此时,它进入 MSC-2 控制的区域,包含 VLR2。位于本地位置的移动台,其本地位置寄存器(HLR)维护着它的 ID 记录、认证信息和当前位置信息(注意,图 8.37 是图 8.1 的一个例子)。

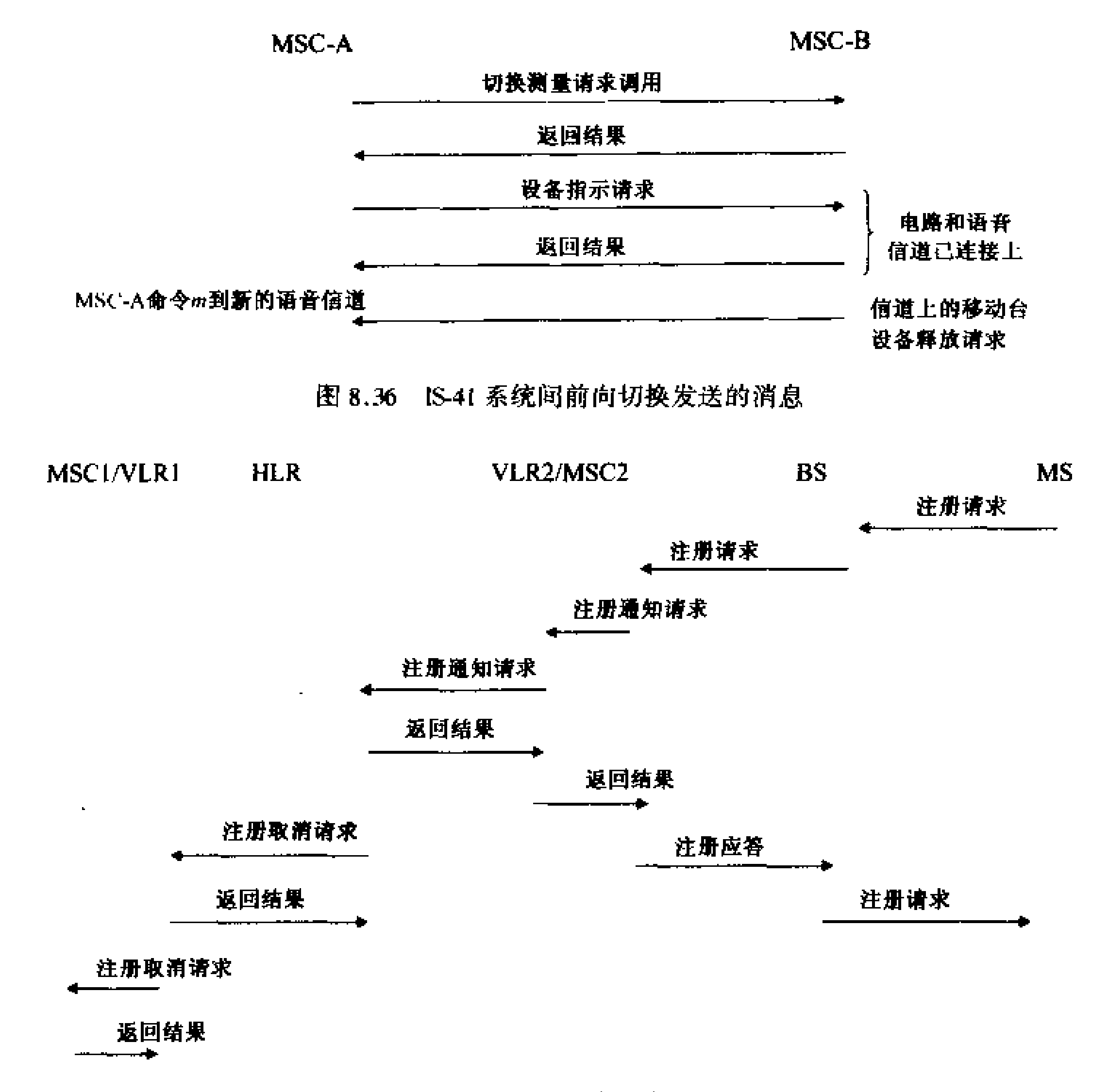


图 8.36　IS-41 系统间前向切换发送的消息

图 8.37　IS-41 位置注册

进入区域 2 的移动台,给基站发送注册消息,然后将该消息转给 MSC 区域。该消息包含移动台识别号 MIN。此时,MSC-2 发送 IS-41 定义的注册 - 通知 INVOKE 消息,包含其 VLR 的 MIN(VIR 实际上可能位于 MSC)。VLR2,通知移动台当前没有被注册,将消息转给移动台的 HLR,使用 MIN 决定 HLR。HLR 消息确认接收到该消息后,将 REGIS.NOT.RETURN RESULT 消息反馈给 VLR2,进入移动台新位置,然后发送注册取消 INVOKE 给 VLR1,将该消息转给它的 MSC。取消移动台在区域 1 内的注册。其他确认消息确认接收到的每一个发送消息,如图 8.37 所示。特别地,注意到每一次注册需要 8 个 SS7 消息,包括在不同的 VLR,MSC 和 HLR 之间发送的确认。在基站与其 MSC 之间,通过空中接口的 4 个其他消息,由注册结果产生。

这里忽略了处理任何差错或意外事件所必须发送的附加消息。假设注册过程顺利而且正确。即使没有其他必须发送的消息,结束注册消息活动也需要大小合适的根据位置区域的大小和其中移动台的密度情况进行的系统容量划分。注册活动影响移动台功率。由于增加了位置区域的尺寸,明显减小了移动台需要注册的边界交叉,所以减小了每个移动台的注册活动。但是在讨论移动台寻呼时将会看到,增加区域尺寸导致在位置区域发送更多的寻呼消息来定位移动台。所以在定位管理和寻呼之间存在折中,在本节开始我们已经提到。我们将使用移动台管理的简单模型定量分析这个折中,粗略决定位置区域的合适尺寸。

现在考虑呼叫传递在其本地漫游,呼叫发起于区域外面空闲的移动终端,这个过程导致移动终端的寻呼,如图 8.38 所示。来自呼叫发起终端的呼叫发起消息(这可能是移动台本身或者有线网络中的固定电话),根据该消息承载的目的移动台号码,传送给最近的 MSC。MSC 使用定位请求 INVOKE(这里简略为 LOCREQ)将该消息转发给目的移动台的 HLR。然后,HLR 发送路由请求 INVOKE(ROUTREQ)给 VLR,被认为是目的移动台注册的最后一个 VLR。VLR 将该消息转发给当前正在使用的 MSC。MSC 给目的呼叫分配一个临时位置目录号(TLDN),路由请求回复(RSP)包含该数字,通过它的 VLR 反馈给 HLR。然后,HLR 发送位置请求回复(RSP)给发起的 MSC,包括移动台的 MIN。发起 MSC 建立一个端到端的语音连接,用在 MSC 为移动台的服务中。然后该 MSC 通过寻呼消息通知它所在区域内所有基站寻呼目的移动台。每一个基站广播寻呼消息给它所在蜂窝中的所有移动终端。空闲的漫游移动台被寻呼,识别其 ID,使用它的随机接入信道回复基站寻呼,这在本节前面已经讨论过。注意到呼叫转接/寻呼过程需要 6 个 SS7 消息,包括从正在服务的 MSC 到每一个区域基站的寻呼消息,每一个基站的广播寻呼消息,以及被寻呼的移动台回复的寻呼消息。对于有 N 个蜂窝的位置区域,每一个寻呼一共有 $8+2N$ 个消息传输给空闲的漫游移动台。随着区域内蜂窝数量的增加和 MSC 覆盖的有效区域的增大,需要更多的寻呼消息。如前所述,这是定位管理的负面影响,增加 MSC 覆盖的区域,减少漫游的移动台的注册数量。所以,根据最小化的消息活动,有一个位置区域尺寸最佳的选择。下面提供一个简单的方法决定最好的位置尺寸的选择。文章提供了一些参考文献,给出了降低定位管理和寻呼复杂度的更一般的情况。

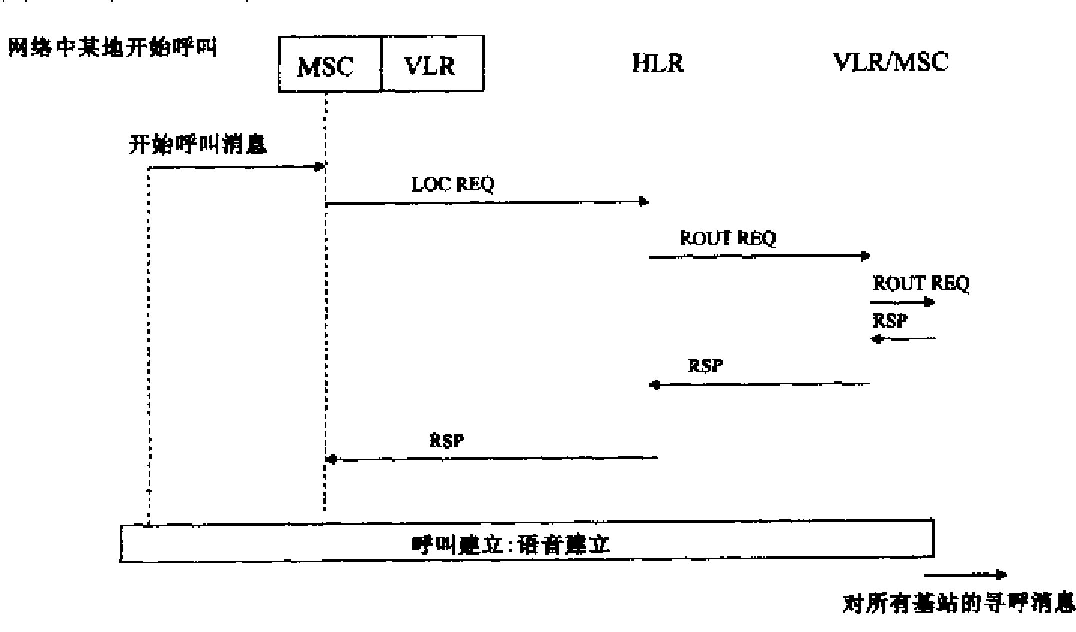


图 8.38 IS-41 呼叫到漫游移动

考虑位置区域 A,包含 N 个蜂窝,每一个面积为 a。区域从它的中心向边界平均扩展 R m,称 R 为区域的半径[1]。圆形或六边形区域的半径为 R。图 8.39 显示了一个正方形区域的例子。令区域内的蜂窝相应的半径为 r m,很明显,必须有 $R=\sqrt{N}r$。该区域有 m 个均匀分布的移动台。移动台的平均速度为 V m/s,各方向分布相同。可以想像的到,平均花费 o(R/V) s 到达边界。典型的移动台经历 o(R/V) = o($V/r\sqrt{N}$)次边界交叉。使用液流模型进行更详细的分析,极其微小的液体颗粒代表移动台,显示穿过尺寸为 S 的区域的平均速度为 $VL/\pi S$, L 为周长或边界长度(Jabbari,1996)。对于半径为 R 的圆形或正方形(参见图 8.39),结果将变成 $2V/\pi R$。对于六边形,结果变为 $2.3V/\pi R$。因此,将使用数值 $2V/\pi R=2V/\pi r\sqrt{N}$ 作为每一个移动台穿过位置区域边界的平均速度。现在,注意到如果每一次穿过边界导致 l 个位置消息被传输,传输位置消息的平均速率当然为 $2lV/\pi r\sqrt{N}$(注意到如图 8.37 所示的 IS-41 中的位置注册情形导致所有类型一共有 $l=12$ 个消息被传输。如果只考虑消息通过空中接口被传输,包括移动终端之间直接传输,则 $l=2$。如果只关心对移动台功率和过程消耗的定位管理的影响,更小的数值将被用于分析。然而,在下面的例子中集中在总的消息传输上)。

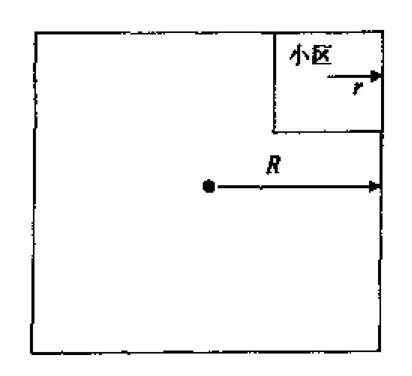


图 8.39　位置区

一个典型的漫游移动终端被呼叫,因此平均每秒被寻呼 λ_p 次。令每一次寻呼需要 p 个消息被传输。对于图 8.38 所示的情况,$p\approx 8+2N$,前面已经提到过。单位时间内传输的寻呼消息的平均数量为 $\lambda_p p=\lambda_p(8+2N)$,单位时间内每一个移动台传输的注册和寻呼消息总数 C 为

$$C=2lV/\pi r\sqrt{N}+\lambda_p(8+2N) \tag{8.6}$$

下面将该数值作为每一个执行定位管理的移动台的开销。当每一个位置区域的蜂窝数量 N 变化时,该数值很明显具有最小值。通过使用令 C 对 N 的导数为零这种通常的技巧,寻找达到最小值时的 N。为了简单起见,假设 N 是连续变量。利用微分得到最佳的 N,定义定位或注册区域的蜂窝数量为

$$N_{opt}=(lV/2\pi r\lambda_p)^{2/3} \tag{8.7}$$

然而,注意到 N 的最佳选择是非常近似的,因为使用的是简化模型,而且是近似分析。考虑使用式(8.7)的一些例子。考虑 3 种情况:宏蜂窝例子、中等尺寸的蜂窝和微蜂窝。令 $l=12$ 个位置消息需要为每一次穿越边界传送,如图 8.37 所示。每个移动台的寻呼速率为每小时 $\lambda_p=$

[1] Jabbari, B. 1996. "Teletraffic aspects of evolving and next-generation wireless communication networks," *IEEE Personal Communications*, 3, 6(December), 4-9.

1 个呼叫到来(下面将考虑把这个速率翻倍。读者可以改变这个速率,检查它们本身对定位区域大小的作用)。对于不同的移动台速率和蜂窝尺寸,N_{opt}的近似值为

1. 宏蜂窝

 (1) $r = 10$ km, $V = 60$ km/h

 (2) $r = 15$ km, $V = 100$ km/h $N_{opt} = 5$ 个蜂窝/区域

2. 中等尺寸的蜂窝(城市环境), $r = 1$ km

 (1)行走情况 $V = 5$ km/h $N_{opt} = 4$

 (2)车载情况 $V = 50$ km/h $N_{opt} = 20$ 或 21

 (如果平均寻呼速率增加到每个移动台每小时 2 个呼叫,则每个定位区域的小区数目将分别下降到 3 个和 13 或 14 个。)

3. 微蜂窝 $r = 100$ m, $V = 5$ km/h $N_{opt} = 20$ 或 21

这些例子证明了式(8.7)通过直觉显示的预期结果:当移动台速度增加,蜂窝尺寸不变时,位置区域增加,以减少边界穿越次数和所需的注册。当蜂窝尺寸减小,移动台速度不变时,位置区域尺寸同样增加。然而,当呼叫达到速率增加时,位置区域尺寸减小。

当然,这些结果是非常近似的,前面已经提到过。它们适用于最简单的可能的流动模型,假设移动台均匀稠密地分布在区域内。只考虑平均速度和平均寻呼速率,更详细的研究和分析参见相关文献。一些研究包括了移动台位置和速度的概率分布。很多减小定位和寻呼开销的不同算法被研究和对比。对定位管理的研究仍然是一个活跃的领域。感兴趣的读者可以参考一些涉及这方面研究和分析的论文和参考文献,包括 Wang(1993)①, Mohan and Jain(1994)②, Bar-Noy *et al*.(1995)③, Akyildiz and Ho(1995)④, Jain and Lin(1995)⑤, Rose and Yates(1995)⑥, Lin(1997)⑦, Akyildiz *et al*.(1998)⑧。

8.5 语音信号处理和编码

到目前为止,在本章中集中讨论了第二代系统的信号消息,以及它们如何通过信号信道进行传输。本节最后考虑这些系统的主要应用,传输语音信号。语音信号的传输速率与有线数

① Wang, J. Z. 1993. "A fully distributed location registration strategy for universal personal communications systems," *IEEE Journal on Selected Areas in Communications*, 11, 6(August), 850 – 860.

② Mohan, S. and R. Jain. 1994. "Two user location strategies for personal communication services," *IEEE Personal Communications*, 1, 1(1st Qtr.), 42 – 50.

③ Bar-Noy, A. *et al*. 1995. "Mobile users: to update or not to update?" *Wireless Networks*, 1, 2(July), 175 – 185.

④ Akyildiz, I. F. and J. S. M. Ho. 1995. "Dynamic mobile user location update for wireless PCS networks," *Wireless Networks*, 1, 2(July), 187 – 196.

⑤ Jain, R. and Y.-B. Lin. 1995. "An auxiliary location strategy employing forwarding pointers to reduce network impacts of PCS," *Wireless Networks*, 1, 2(July), 197 – 210.

⑥ Rose, C. and R. Yates. 1995. "Minimizing the average cost of paging under delay constraints," *Wireless Networks*, 1, 2(July), 211 – 219.

⑦ Lin, Y.-B. 1997. "Reducing location update costs in a PCS network," *IEEE/ACM Transactions on Networking*, 5, 1(February), 25 – 33.

⑧ Akyildiz, I. F. *et al*. 1998. "Mobility management in current and future communication networks," *IEEE Network*, 12, 4(July/August), 39 – 49.

字电话系统使用的速率相比非常低。这是必须的,因为在无线蜂窝系统中可用带宽相对较小。恶劣的传输环境包括衰落和来自其他移动终端的干扰,需要通过编码进行很强的差错保护。所以,通过空中接口传输语音信号需要两个步骤。首先,语音信号必须被有效地压缩,以降低传输所需的比特速率;然后,使用编码技术提供差错保护。这两个步骤使得语音信号可以被正在接收的用户所接收。

语音信号压缩已经被研究很多年了。很多技术被提出、调研、标准化和实现。早期的数字电话系统中的64 Kbps PCM语音传输,是通过对3.2 kHz带宽受限的语音每秒进行8000次采样获得的,然后每次采样使用8比特提供电话应用可接受的速率质量。该技术依赖于Nyquist采样定理的应用。该定理指出,通过对带宽受限的信号以两倍带宽的速率进行采样,可以无差错地恢复原始信息(Schwartz,1990)。使用8比特或256个信号幅度等级的量化过程(Schwartz,1990)[①],引入了一些误差,称为量化噪声。差分PCM将当前的语音信号采样,再与以前采样值加权和进行比较,然后将差异信号进行量化和传输,可以有效地降低电话语音信号传输速率至32 Kbps。降低传输速率需要更多的精密和复杂的信号处理技术。

为什么可以压缩语音信号的传输速率呢?我们都知道语音通信中存在着相当大的冗余度。例如,听者经常可以将说话者丢失的文字或一部分组合起来,通过它们所在的上下文。以前每秒8000次采样的PCM实验显示了本地听者一般可以理解量化到每个采样1比特的语音信号(但是,噪声却相当高)。所以,目的是在保证合理的信号质量下,尽可能地降低冗余度。在数字语音信号通过恶劣的介质例如无线蜂窝系统进行传输时,存在附加的折中。前面已经看到:当信号冗余度降低时,它对于衰落、噪声和空中接口引入的干扰变大了。所以,我们需要对传输的数字信号进行编码,提高所需的传输比特速率,有效地降低冗余度压缩。使用特定的信号压缩技术变得至关重要。

第二代蜂窝系统将语音信号压缩为未编码的数据率,该数据率IS-136中的8 Kbps到GSM中的13 Kbps进行变化,它们都采用了一种称为线性预测编码(LPC)的压缩技术。LPC本质上起始于对人类语言的演变过程建立模型,作为激励波形来驱动一组滤波器(Atal和Hanauer,1971[②];Jayant,1992[③];Rabiner和Hwang,1993[④])。LPC语言模型如图8.40所示。激励波形和长期滤波器建立起肺部和声带产生的语言声音的模型;而短期滤波器代表声音束。如图8.40所示的LPC模型,其中激励波形是周期性脉冲与类似噪声信号的混合。术语长期用于描述图8.40中的第1个滤波器,粗略涉及到声带振动引起的语音声音的音调周期。音调周期一般在3~15 ms数量级的范围内(Rabiner and Hwang,1993)。短期暗示时间则在1 ms数量级上。

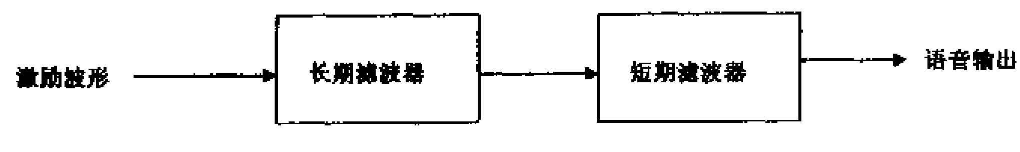


图8.40 LPC语音模型

① Schwartz,M. 1990. *Information Transmission, Modulation, and Noise*, 4th edn, New York, McGraw-Hill.

② Atal, B. S. and S. L. Hanauer. 1971. "Speech analysis and synthesis by linear prediction of the speech wave," *Journal of the Acoustic Society of America*, 50, 2 (August), 637-655.

③ Jayant, N. 1992. "Signal compression: technology targets and research directions," *IEEE Journal on Selected Areas in Communications*, 10, 5(June), 796-818.

④ Rabiner, L. and B.-H. Hwang. 1993. *Fundamentals of Speech Recognition*, Englewood Cliffs, NJ, Prentice-Hall.

为什么称之为“线性预测编码”？LPC 语言模型的输出在给定采样时间内基于以前采样值的加权和，再加上根据采样输入得到的新数值。具体而言，集中在图 8.40 所示的短期滤波器。该滤波器的简单框图如图 8.41 所示。采样间隔 n 时的输出 $\hat{s}(n)$ 通过将采样输入 $e(n)$ 加到线性预测器上得到，如图 8.41 所示。线性预测器模型可以简单地描述为

$$\hat{s}(n) = \sum_{i=1}^{m} a_i \hat{s}(n-i) + e(n) \tag{8.8}$$

当前输出值基于 m 个以前的采样值，相邻间隔 125 μs，前面已经提到过。a_i 是自适应加权因子。对式(8.8)进行 z 变换，读者可以得到传输函数 $H(z) = S(z)/E(z)$ 为

$$H(z) = 1\bigg/\left(1 - \sum_{i=1}^{m} a_i z^{-i}\right) \tag{8.9}$$

该传输函数正好是全极点线性滤波器。采样数 m 的典型值在 10 左右，这意味着滤波器工作在相对“短”的 1 ms 时间内，前面已经提到过。长期滤波器模型通常也是全极点滤波器，但是仅有一两个极点。具体而言，它由单极点滤波器表示，其传输函数可以显示表示为

$$H(z) = 1/(1 - bz^{-p}) \tag{8.10}$$

参数 p 为音调周期与采样时间的乘积，被周期性地调整和测量。所以，它在 20 到 120 左右范围变化，对应于实际时间 2.5 ms 到 15 ms，如前所述。参数 b 是加权因子。

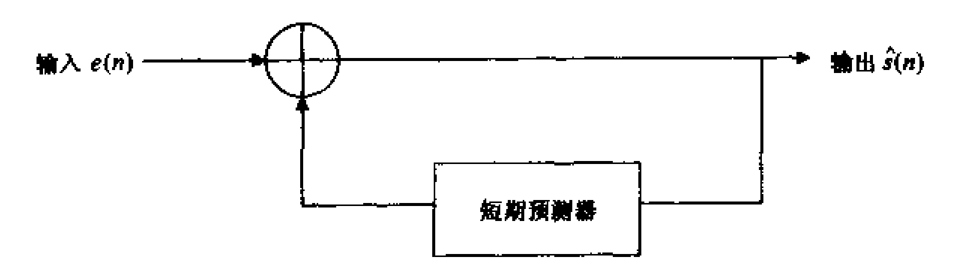


图 8.41 短期滤波器：线性预测

用于压缩语言的基本 LPC 系统，使用图 8.40 和图 8.41 所示的 LPC 语言产生模型，如图 8.42 中的框图所示，也可参阅 Steele，1992：227。系统将模型输出与实际语言采样值进行比较，通过周期性调整激励和滤波器参数，试图减小差异信号，它显示了语言产生端的编码器[如图 8.42(a)所示]和接收端的译码器[如图 8.42(b)所示]。混合系统通常称为语言编译码器。首先考虑编码器。量化的输入语言采样值，标记为 $s(n)$，每 125 μs 产生一次。这些和语言模型输出 $\hat{s}(n)$ 之间的差异 $\epsilon(n)$，通过调整激励发生器和滤波器参数被最小化[对几个样本求最小均方误差，也就是最小化 $\sum_n \epsilon^2(n)$]。然后，结果参数在采样间隔被传输给接收系统。接收系统，即译码器，则执行相反的过程，使用接收参数调整激励发生器和滤波器。下面将会看到，周期性发送参数，而非语音采样或者它们的差异，会导致传输速率大幅降低。第二代系统中所使用的编译码器将运用该系统的变化，如前所述。

特别地，GSM 系统使用了称为 LPC-RPE(LPC with regular pulse excitation)策略。IS-136 和 IS-95 使用 CELP(code-excited linear predictive coding)策略作为语音编译码器。现在描述这些不同的

编码-译码技术。首先,考虑 CELP 技术(Schroeder 和 Atal,1985[1];Chan 等,1992[2];Steele,1992)。基本的 CELP 策略的框图如图 8.43 所示(图中未显示的是加权滤波器,它通常用于加权输入信号采样和/或差错采样)。注意到系统与图 8.42 所示的 LPC 编译码器看起来非常类似,主要差别在于图 8.43 所示的激励发生器通过选择一组高斯码字给定。参数 G 代表调整选定码字幅度的增益参数。图 8.43 所示的长期滤波器通常被设计为式(8.10)所示的单极点滤波器,传输函数为 $1/(1-bz^{-p})$。如前所述,参数 p 代表伪周期性语音声音的估计音调周期。参数 b 为可调整的加权因子。随机码字和长期滤波器的混合用于建立声带激励模型,在讨论 LPC 时已经提到过。短期滤波器还用于建立声音束模型。

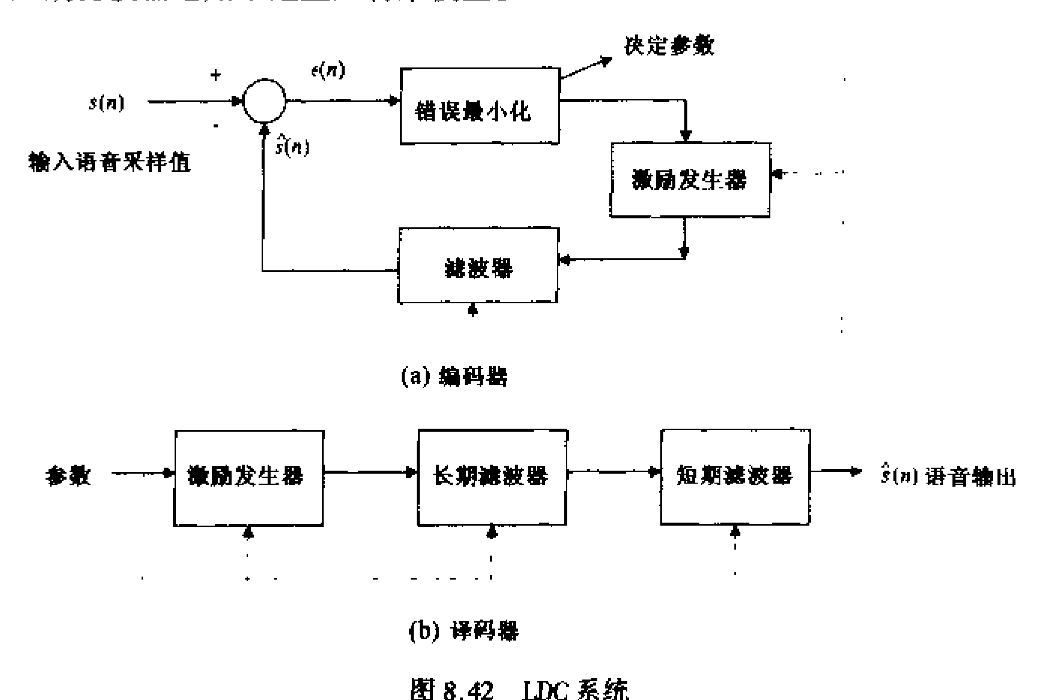


图 8.42 LDC 系统

CELP 使用向量量化(Rabiner 和 Hwang,1993):这是一个在收集到很多量化差错构成的向量以后进行周期性参数调整的技术。向量量化提供压缩优势。例如,每 5 ms 调整一次,125 μs 进行一次语言采样,所以每 5 ms 产生 40 个采样值构成的向量。激励码字集包含适当数量的随机码字,每一个含有 40 个量化元素的向量。滤波器参数、增益和码字均方误差最小化如图 8.43 所示,用于决定每个间隔的最好参数和码字。正是这些参数通过空中接口语音信道被发送,用于译码器重建语言信号,如图 8.43(b)所示。如前所述,传输这些参数而非实际语言信号的低比特表示,导致语音信道的比特速率在很大程度上被压缩。考虑一个由 IS-136 编译码器产生的简单例子。短期滤波器是一个 10 极点全极点滤波器,传输函数为 $1/(1-\sum_{i}^{10}a_i z^{-i})$。该滤波器有 10 个待定系数。这些系数每 20 ms 发送一次,每个系数用 4 比特就可以精确表示,所以,编码器每 20 ms 仅需要 40 比特就可以表示它们。令其他参数为音调周期 p、加权系数 b、增益 G,被选码字每一个用 7 比特表示,每 5 ms 收集一次,这对应于每 5 ms 28 比特。每 20 ms

① Schroeder, M. R. and B. S. Atal. 1985. "Code-excited linear prediction (CELP): High quality speech at very low bit rates," Proc. IEEE Int. Conf. Acoust. Speech, Signal Process., March, 937 – 950.

② Chen, J.-H. *et al*. 1992. "A low-delay CELP coder for the CCITT 16 kb/s Speech Coding Standard," *IEEE Journal on Selected Areas in Communications*, 10, 5 (June), 830 – 849.

发送的总比特数为 152 个，比特速率为 7600 bps(在下面的 IS-136 中，我们可以看到通过语音信道每 20 ms 发送 159 比特，则总的比特速率为 7950 bps)。所以，使用向量量化 CELP，比特速率得到足够的降低是可能的。介绍完一般的语音压缩技术以后，下面具体讨论在每一个第二代蜂窝系统中使用的语音压缩/编码技术。

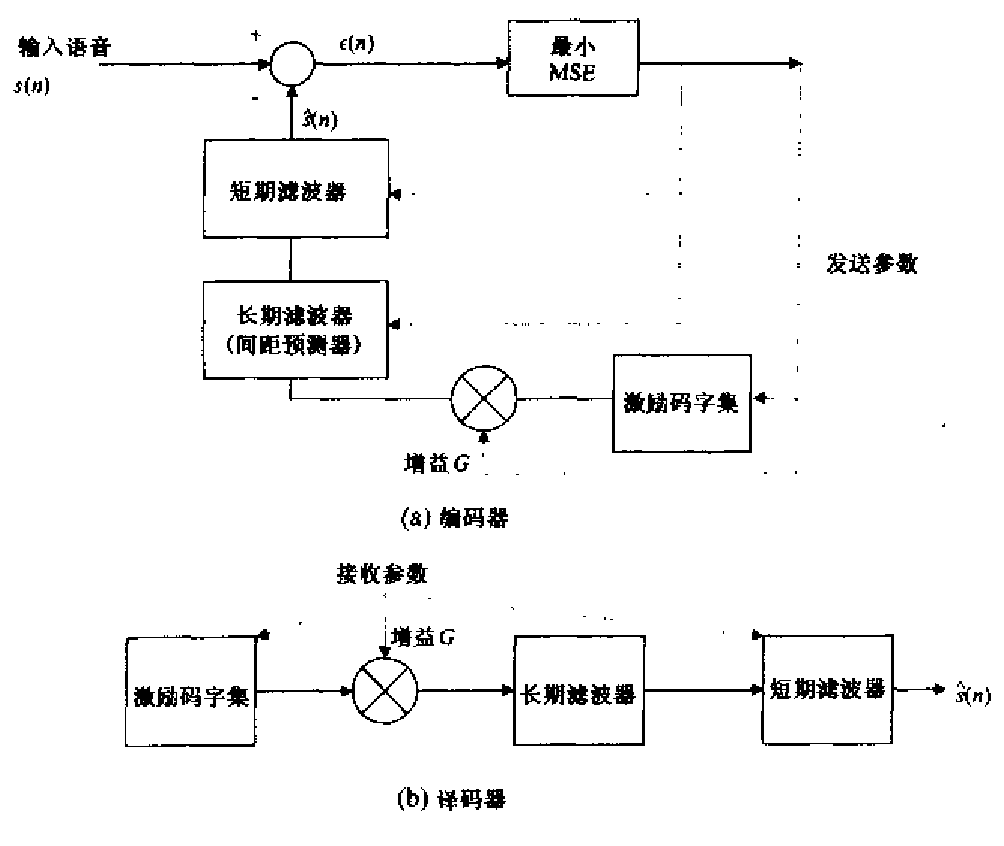


图 8.43 CELP 系统

IS-136

北美第二代蜂窝系统 IS-136 语音压缩系统使用 CELP 的变形 VSELP(vector-sum excited linear predictive coding)。在此系统中，模拟语言以每秒 8000 次采样的速率被采样，前面已经多次提到过，所以，它的采样间隔为 125 μs。然而，采样值被量化为 13 比特/采样，而非有线电话中 8 比特/采样。VSELP 框图如图 8.44 所示(TIA, 1992)[①]。注意到图 8.44 所示的 VSELP 编译码器与图 8.43 所示的基本 CELP 编译码器非常类似，其主要差别在于 VSELP 使用两个随机码字集合，而 CELP 只使用一个。长期滤波器被式(8.10)所示的单极点滤波器表示，其音调参数由图 8.44 中的参数 L 给定。图 8.44 中的增益参数 β 与式(8.10)中的参数 b 对应。L 数值在 20 到 146 范围内变化，是采样间隔 125 μs 的倍数，对应于 127 种可能的数值，所以它可以用 7 比特表示。编码器的加权合成滤波器 $H(z)$和滤波器 $A(z)$都与式(8.9)所示的全极点传输函数相对应，具有 $m = 10$ 个加权参数，记为 α_i，$i = 1 \sim 10$，如图 8.44 所示。信号采样值也被加权滤波器 $W(z)$加权，该滤波器同时具有零点和极点，$W(z)$的分子(零点)和分母(极点)具有相同

① TIA, 1992. "EIA/TIA Interim standard, cellular system dual-mode mobile station-base station compatibility standard," 15-54-B, Telecommunication Industry Association.

的参数 α_i。每个码字集合中的码字由 40 个元素组成的随机向量给出，由从每一个码字集合中选出的 7 个 40 元素的随机基向量定义。每一个码字集合一共有 $2^7 = 128$ 种可能的码字，可以构成适当的激励向量。

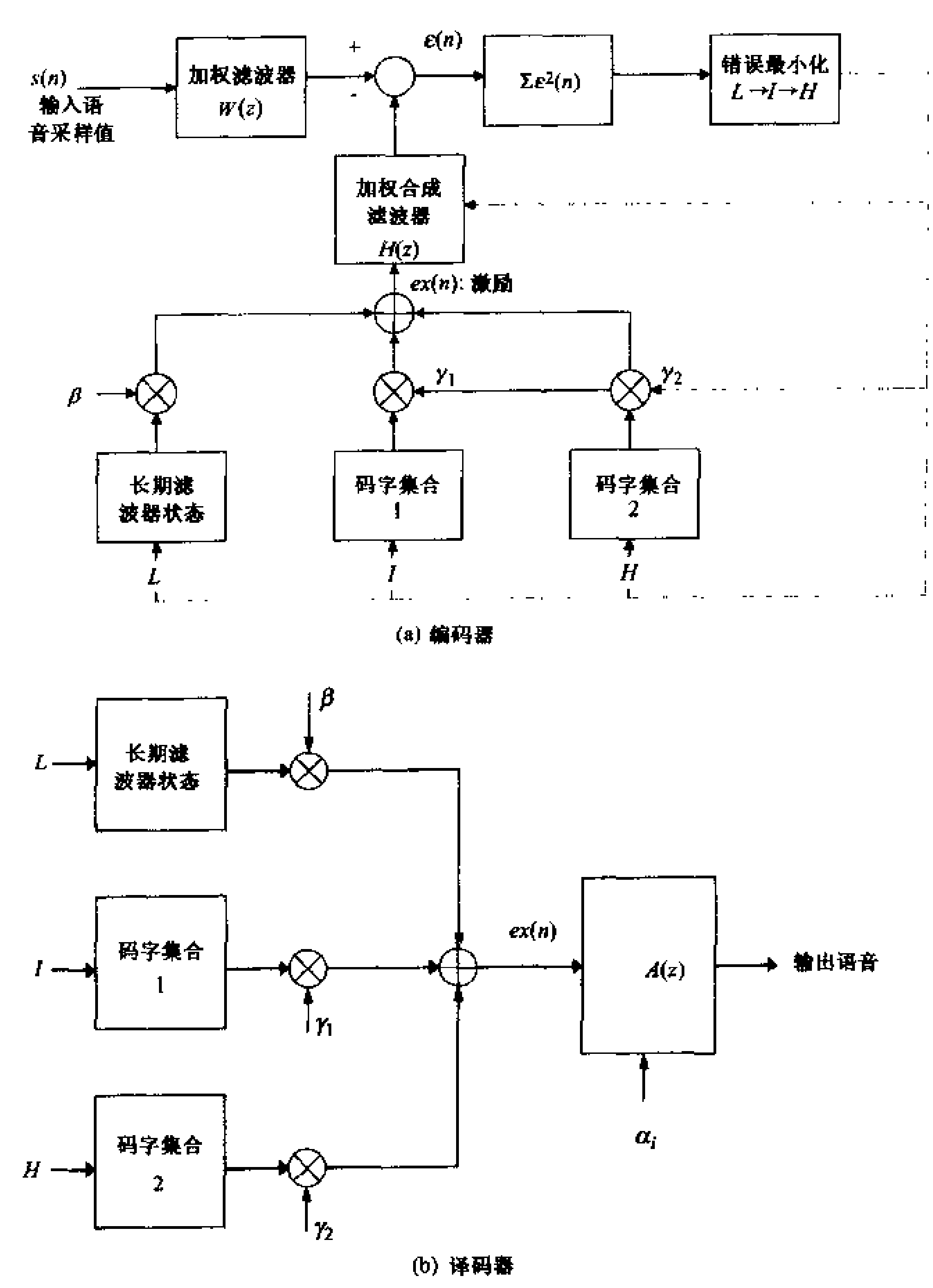


图 8.44　IS-136 编译码器：VSELP 系统

编码器工作过程如下：差错项 $\varepsilon(n)$ 被平方，每 5 ms 间隔累加一次，相应于 40 个信号采样。然后，在不同的码字和参数下被最小化，首先将长期滤波器作为惟一激励进行最小化，然后转到码字，如图 8.44 中的参数 I 和 H。用于最小化平方误差的参数和码字，通过基站和移动台之间的空中接口传输给图 8.44(b)所示的接收译码器。这里值得强调的是，压缩的信号采样值或差异值没有通过空中接口传输，如同 PCM 和 DPCM 一样。自适应变化的 CELP 参数

被传输到译码器。图 8.44(b)所示的译码器,如同图 8.43(b)中 CELP 系统的译码器一样,复制图 8.44(a)所示的编码器语言模型。参数和码字每隔 20 ms 或每隔 5 ms 被传输一次。为什么选择 20 ms 间隔发送 VSELP 参数呢? 回忆在第 2 章中讨论的 IS-136 系统,每 40 ms 帧有两个时隙,每一个时隙载有 260 个数据比特,用于提供全速率语音传输(也如图 8.9 所示)。每 20 ms 间隔传输 VSELP 参数,每个 IS-136 帧提供两次编译码器自适应。这些参数作为 159 比特每 20 ms 被发送一次,如前所述,大致相当于 7.95 Kbps 语音速率,如前面提到的 IS-136 压缩比特速率。下面描述的编码导致语音编码比特速率为 13 Kbps。159 比特在不同参数和码字之间被划分如下:3 个增益 β, γ_1 和 γ_2,每 20 ms 间隔总共被分配 8 比特。每 5 ms 子间隔需要 7 比特传输参数 L,前面已经提到过,20 ms 总共需要 28 比特。每 20 ms 38 比特用于传输 10 个滤波器参数。最后,每 20 ms 帧的能量也被传输,使用 5 比特。每 20 ms 总共 159 比特,这在开始已经提到。

这 159 比特在传输之间必须被保护,从而避免可能的信道差错。3 种保护机制被使用。前向差错控制,某些比特使用循环冗余编码 CRC 和卷积编码,与剩余未控制比特一起,用于将 159 比特原始参数转变为 260 个数据字段比特,每 20 ms 被传输一次。然后,这 260 比特在 40 ms IS-136 帧的两个时隙中交织。图 8.45 描述了整个差错控制过程。语音通信的差错控制过程与图 8.12 所示的 IS-136 控制信息的差错控制过程进行比较,详细描述如图 8.45 所示,注意到 159 比特首先被划分成两组,第 1 组 77 比特,第 2 组 82 比特,如图所示。第 1 组比特仅被差错控制机制保护。例如,码字参数中没有保护的比特都作为第 2 组被发送。L 参数都作为第 1 组参数发送,其他参数作为混合组比特发送。第 1 组 77 比特中有 12 比特被选做最重要的比特。这 12 比特选自能量参数和 10 个滤波器参数中的 5 个。这 12 比特首先附加 7 个 CRC 比特,提供一些前向差错控制,如图 8.45 所示。经过 CRC 编码后的序列和剩余第 1 组比特被传送给效率为 1/2 的卷积码编码器,附加 5 个尾比特,形成 89 比特卷积码。卷积码编码器生成的 178 比特附加第 2 组 82 个未保护比特,如图 8.45 所示,每 20 ms 形成 260 比特的输出序列。

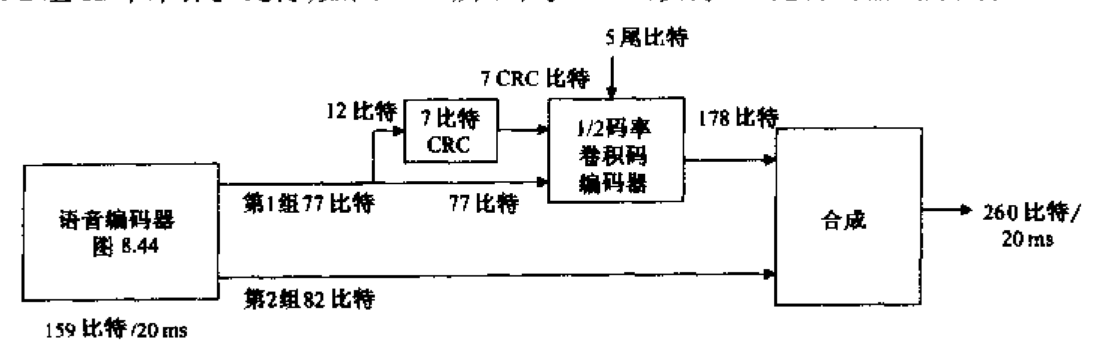


图 8.45 IS-136 语音编码器编码

这 260 比特在两个数据通信时隙中被交织,进行进一步保护,前面已经提到过(参见图 8.9)。交织过程如图 8.46 所示。连续编码的 260 比特序列被标记为 a, b, c, d, e, f, g…每一个序列中有一半出现在分配给特定的语音信道的两个连续时隙中的每一个上(回想 IS-136 中 6 个时隙持续 40 ms,每帧 2 个时隙被分配到 3 个信道中的每一个,如图 8.9 所示)。图 8.46 显示了 260 个 a 比特中的后一半加上 260 个 b 比特中的前一半出现在一个时隙中的情况;b 比特中的后一半和 c 比特中的前一半分配给同一信道中的下一个时隙,依次类推。每一个语音用户的这 260 个编码比特在两个连续帧中发送,跨越 3 个时隙间隔。

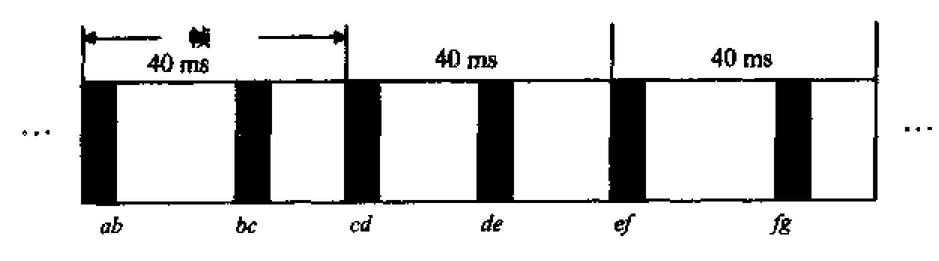


图 8.46　IS-136 交织的效果

IS-95

IS-95 使用的语音压缩技术与前面讨论的 IS-136 中的语音压缩技术类似，除了只使用一个码字集合以外(Garg,2000)[①]。它使用了 CELP 型技术，如图 8.43 所示。语言采样也是以 125 μs 间隔进行的。在这种情况下，至少量化为 14 比特/采样，然后与 LPC 型滤波器的输出进行比较，其输入是一个从差错码字集合中选择的随机向量。每 20 ms 执行一次最小均方误差决定，如同 IS-136 系统一样，最小化均方误差的 CELP 系统参数被传输给接收系统的译码器，如图 8.43 所示，这在讨论 IS-136 中的 VSELP 系统时描述过。

在 8.3 节中提到，对于 IS-95，可以使用很多种语音比特速率。我们还是只关注全速率的实现。考虑图 8.18 和图 8.19 所示的全速率 IS-95 系统反向和前向业务信道框图。它们表明每 20 ms172 比特，被定义为帧间隔，提供这些系统的输入。正是这些输入比特与在编码和调制以后通过空中接口传输给译码器的 CELP 参数相对应。前向差错控制、卷积编码和块交织功能的描绘在这些图中，在 8.3 节中进行了讨论，提供了与前面所讨论的 IS-136 系统相对应的语音保护功能，如图 8.45 所示。

更准确地说，，每 20 ms 帧 171 比特被 CELP 译码器输出，用于 IS-95 语言比特速率 8.55 Kbps。回忆 8.3 节中的表 8.3，IS-95 全速率业务信道可以用于传输主要业务和信号业务。仅考虑信道载有主要业务的情况。根据表 8.3，由于有一个比特用于指定这种情况，每 20 ms 帧实际发送 171 比特。这 171 比特分布在不同的系统参数中，如同 IS-136 VSELP 译码器的情况。IS-95 CELP 译码器使用单极点长期滤波器。它的音调参数 L 每 5 ms 被更新一次，如同 IS-136，或者每帧更新 4 次，每次更新传输 28 比特中的 7 比特。每 5 ms 子间隔，3 个比特或者每帧总共 12 比特用于传输长期滤波器的增益系数。每帧传输 40 比特短期滤波器系数。激励码字和增益 G 每帧更新 8 次(参见图 8.43)，7 比特和 3 比特分别用于代表每次更新的参数。所以，码字集合传输需求为每20 ms 帧 80 比特。需要传输的这些 CELP 参数总共为每帧 160 比特。然而，为了进一步控制差错，如图 8.18 和图 8.19 所示，每帧有 11 个奇偶比特被计算并附加在 160 比特后面，得到前面提到的总共 171 比特。

GSM

本节最后讨论第二代数字蜂窝系统 GSM 中的语音信号压缩和编码。如前所述，GSM 语音压缩策略称为带有规则脉冲激励的线性预测编码(LPC-RPE)。其中，LPC 是这个策略的基础，如同它用于前面讨论的两种压缩过程一样。它紧随图 8.42 所示的基本 LPC 形式。GSM 技术和其他两种蜂窝系统基于 CELP 技术的差别在于激励发生器的选择。GSM 使用幅度可变的规

① Garg, V. K. 2000. *IS-95 CDMA and cdma2000*, Upper Saddle River, NJ, Prentice-Hall PTR.

则脉冲，提供滤波器激励，而非其他系统所采用的 CELP 随机码字集合的方法。选择这些脉冲的幅度使得均方误差最小，如图 8.42(a)所示。GSM 系统采用的方法由于其他两种策略，提供小一点的压缩效率：每 20 ms 间隔 260 比特由语言编码器产生，实现 13 Kbps 有效语言传输速率。最近，增强的全速率和半速率 CELP 型编码器被 GSM 系统标准化。半速率系统工作在 5.6 Kbps，为 VSELP 类型(Garg 和 Wilkes，1999)①。

LPC-RPE 策略的详细讨论参见 Steele 的著作(1992：Ch.3 和 Ch.8)。所用的长期预测滤波器是单极点类型。延时和增益参数使用最小均方误差估计，每 5 ms 被确定一次，7 比特和 2 比特分别用于代表延时和增益的参数。这些占有 260 比特中的 36 比特，每 20 ms 被组合一次。对短期滤波器的 8 个系数进行的估计，每 20 ms 占有另外 36 比特。剩余 188 比特，每 5 ms 47 比特，代表定义合适的激励信号的参数(Garg 和 Wilkes，1999；Steele，1992：Ch.8)。

紧随 LPC-RPE 语言编码过程的差错保护如图 8.47 所示(Good man，1997；Steele，1992：Ch.8)。在差错保护过程后进行交织。该保护过程与图 8.6 到图 8.8 所示的 GSM 中控制信道的保护过程进行了比较。如前面所述的 IS-136 系统的情况(参见图 8.45)，260 比特语言编码器的输出每 20 ms 出现一次，被分成两组。其中，182 比特被分成第 1 组，称为重要比特；78 比特被指定为第 2 组，称为非重要比特。差错保护不用于非重要比特。第 1 组比特进一步被分成两组，如同 IS-136 系统一样：182 个比特中的 50 个通过 3 个奇偶校验比特首先被保护，检测错误但不纠正错误。然后这 53 个比特和剩余 132 个比特一起被码率为 1/2 的卷积码编码器编码，附加 4 个尾比特。整个框图如图 8.47 所示。

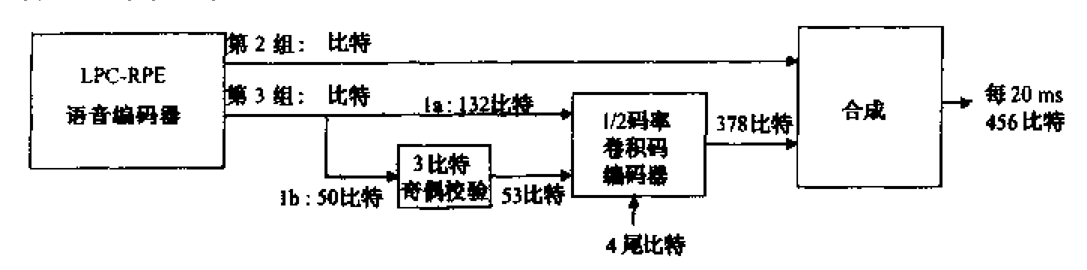


图 8.47 GSM 语音信号保护编码

每 20 ms 出现在卷积码编码器输出端的 378 比特加上未被保护的 78 比特，每 20 ms 456 比特，提供速率为 22.8 Kbps 的语音通信。回忆 8.1 节，这正是 GSM 系统采用的语音通信比特速率(如图 8.2 所示，114 个数据比特在每一个通信时隙被传输。全速率用户每帧占用这样一个时隙，4.615 ms 间隔重复一次。13 帧中有一帧被抽取出来用于控制的目的)。交织操作在前面提到的保护编码操作后被执行，提供进一步的信号保护。注意到 456 是 57 的 8 倍。在交织操作中，在两个 456 比特长的语言块中每一个有 57 比特，占用 40 ms 时间，组合在一起在 8 个连续帧中填充时隙(Goodman，1997；Garg 和 Wilkes，1999；Steele，1992；Ch8)。

习题

8.1 图 8.3 显示了如何建立 GSM 呼叫，使用文中和表 8.1 中所描述的各种控制信道。建立呼叫所需的各个步骤如图 8.3 所示，按照时间在图 8.4 和图 8.5 的时间框图中描述。为

① Garg，V. K. and J. E. Wilkes. 1999. *Principles and Applications of GSM*，Upper Saddle River，NJ，Prentice-Hall PTR.

了这个目的，准备两组复帧时间框图，一个在另一个的下面。一个代表试图建立连接的给定移动台的传输；另一个代表对应于移动台通过上行链路传输信息的基站，其定时所用的帧。这两个框图指示出，当呼叫建立时，各种控制信号传输给移动台或者从移动台传输，如同它们出现在每一帧中一样。

8.2 比较表 8.2 所示的在 IS-136 中的控制信道和表 8.1 所示的在 GSM 中的控制信道。比较考虑特殊控制信道的功能，以及它在时隙/帧结构中的实现。哪些控制信道类似，哪些不同?

8.3 (a) 提供如图 8.3 所示的框图，显示 IS-136 系统如何建立呼叫。这需要增加图 8.15 所示的接入过程，包括移动台通电以及获取所需的同步和接入信息。

(b) 对于 IS-136 系统重复回答习题 8.1，将(a)中的框图转化为合适的复帧时间框图。

(c) 比较 GSM 系统和 IS-136 系统建立呼叫的过程。

8.4 考虑如图 8.18 和图 8.19 所示的 IS-95 业务信道框图。

(a) 解释为什么图 8.19 代表 QPSK 调制系统；图 8.18 代表 OQPSK 调制系统。

(b) 显示出现在图顶端的各种数据比特率如 9.2 Kbps、28.8 Kbps、307.2 Kbps 和 19.2 Kbps 是如何获得的。

(c) 和上行方向相比，为什么下行传输方向需要较小的差错校正能力?

(d) 解释为什么块交织可减小突发错误的影响。

(e) 显示对图 8.19 所示的 19.2 Kbps 数据序列进行凿孔是如何导致 800 bps 功率控制信号的。

8.5 (a) 显示式(8.4)给出的 Walsh 码 W_8，找到 Walsh 码 W_{16}。

(b) 证明式(8.5)所示的(a)中的两个 Walsh 码的正交关系。通过观察，显示这两个码都具有一半 0、一半 1，所有行分别有 4 个和 8 个元素不同。

8.6 (a) 解释图 8.23 所示的细节，集中在每一个块的功能，以及显示如何确定不同的比特速率。

(b) 参考图 8.25，显示 IS-95 移动台通电后发起呼叫的过程，解释移动台从导频、同步和寻呼信道中得到了什么信息。

(c) 将图 8.25 和图 8.26 所示的 IS-95 系统中的呼叫开始和定向过程与 GSM 和 IS-136 (D-AMPS)进行比较。

8.7 图 8.27 给出了一个 dim-and-burst 帧的帧结构的例子。为表 8.3 中其他例子重复这个图。

8.8 图 8.29 给出了 IS-95 中复帧信号的一个例子。显示需要使用 blank-and-burst 信号传输 2016 比特长信号消息的帧序列。表示每一帧的比特分配，如图 8.29(b)所示。为 dim-and-burst 帧重复这个过程。

8.9 用自己的语言解释硬切换和软切换的差别。为什么软切换为 CDMA 系统提供了性能改善(参见第 6 章)?

8.10 图 8.36 描绘了当系统之间的切换过程被执行时需要被传输的消息交换。参考书中系统之间的切换反馈和切换到第三方的过程。为每一个其他过程提供可能的消息交换框图。

8.11 令 r 为蜂窝的半径，单位为 m；R 为包含 N 个蜂窝的位置区域的半径，单位为 m。证明为什么平均移动速度为 V m/s 的典型移动台平均每秒 $o(V/r\sqrt{N})$ 次越过位置边界。

8.12 (a) 证明对于文中折中讨论的简单位置区域,式(8.7)给出了每个位置区域最佳蜂窝数的估计。

(b) 检验文中所选例子中给出的每个位置区域的蜂窝数值,并在不同的蜂窝大小、移动速度和寻呼速率的情况下,重新计算。

8.13 (a) 证明对于图 8.40 所示的 LPC 语音模型中的短期滤波器,图 8.41 所示的线性预测模型有式(8.8)的形式。

(b) 证明(a)中滤波器传输函数的 z 变换形式为式(8.9)。解释这个陈述:对此滤波器使用 $m=10$ 个样值,该滤波器操作时间大约为 1 ms。

(c) 解释在图 8.43 所示的 CELP 系统中,若每 5 ms 进行一次调整的话,为什么会生成 40 个元素的向量。

(d) 完成文中描述的 CELP 编译码例子的计算,并给出传输比特率为 7600 bps 时的结果。

8.14 (a) 考虑图 8.44 所示的 IS-136 VSELP 语音编译码器,完成文中提到的计算并进行解释,首先,如何获得 7.95 Kbps 未编码语音比特率,然后,在图 8.45 所示的编码后比特率为 13 Kbps 。

(b) 对文中讨论的 GSM 语音编码器,在如图 8.47 所示的有保护编码的情况下,重复问题(a)。首先计算未编码比特率,然后解释实际的传输比特率为 22.8 Kbps。

第9章　性能分析：准入控制和切换

在本书的前面几章里，已经明确讨论了一个蜂窝系统容量的概念。具体说来，在第2章里利用著名的厄兰B公式把呼叫阻塞的概率同流量强度（呼叫到达速率乘以呼叫持续时间）和容量联系起来。在这里，容量是用信道的个数表示的，每个信道可以处理一个呼叫。在第3章里，介绍了如何引入"蜂窝"的概念来增加整个系统的容量，也就是可以处理的同时发生的呼叫的数量。在第4章进一步介绍了在中等流量强度的情况下，动态信道的分配在理论上可以用来增加系统的容量。在第6章，比较了IS-136，IS-95和GSM这3个第二代数字蜂窝系统的容量。在那一章里，还指出了当移动台在一个特定系统中移动的时候，必须切换到相邻的蜂窝。特别地，指出基于CDMA的IS-95通过软切换在性能上有所提高。在第8章里通过分析实际的切换过程来进一步介绍切换。

在本章里，通过描述蜂窝准入控制来量化对切换过程的讨论：在一个小区里，只有容量还有富余的时候，新的呼叫请求才被接受，否则就会发生呼叫阻塞。向一个新的小区切换也会用同样的方式来处理，在这种情况下，它们的阻塞概率又称为切换掉话，同新业务的阻塞率是一样的，或者小区之间的切换将被赋予某种优先级来降低掉话的概率。如果可能的话，后一种方法更被看好，因为人们更不希望在通话的过程中掉线。尽管如此，如果切换优先级被引入，将存在一个折中：在容量一定的情况下，降低切换掉话的概率就会导致新业务的阻塞率上升。如果新业务的阻塞率增加得不很大，而切换的掉话率却大幅下降的话，很明显采取这样的技术有优势。在文献里提出了大量处理切换的流程，要对它们中的一部分进行描述和比较。在同样阻塞率的情况下，如果一种方案可以获得更低的切换掉话概率，那么它就比另一个好。反过来说，当切换掉话概率限制在一个特定数值以下时，获得较低阻塞概率的方案更好。

一个常见的切换优先级方案是为新的切换预留一定数量的信道。预留信道的数量要与流量强度相适应，满足一定的切换掉话和阻塞概率的比值。这种优先级过程常称为保护信道方法，原因很简单。在文献中还提出了一种方案：相邻小区定期地相互通知各自信道的利用情况。通过提前预测，某个小区就可以估计出一个在相邻小区取消服务后接下来进行切换的呼叫数量。如果可能切换的呼叫数量超过一个阈值，最好先拒绝接入新的呼叫。计算表明，这种过程优于保护信道方案，但是后者不需要相邻小区之间定期的通信。文献中还提出了其他的一些过程。一个简单的方案是如果开始没有空闲信道，那么就将切换业务缓存起来，直到某个最大时间。切换掉话率的确降低了，代价是在连续的服务中产生了延时。如果延时不是太大的话，对通话的参与者来说也许可以接受。

对各种方案的学习和比较需要开发和分析蜂窝系统的性能模型。从9.1节开始，我们开始研究简单模型独立于小区形状的呼叫持续期和切换次数。在9.2节中，研究了一维的小区结构的性能问题。在分析的过程中，很多移动模型被引入和加以比较。它们被用来比较上面提到的接入控制方面的主题。在9.3节，把分析拓展到更一般、也更复杂的二维小区模型。这些越来越复杂的模型，补充和拓展了前些章节讨论过的有限的性能建模和分析。和前面的章

节一样,假设一个系统具有均匀、同分布的移动密度和流量条件。通过本章的学习,应该对蜂窝系统的一般指标有一个更好的了解。

9.1 关于性能概念的概述

在这一节里,首先关注一个小区里呼叫持续期和移动驻留时间的简单概率模型。在这些模型里,准备计算小区里一个信道的持续时间长度,还有切换发生的概率。然后,用简单的流量分析把小区里切换流量强度和新业务的流量强度联系起来(显然,切换业务必须像新业务一样开始)。我们提供了不同小区尺寸的例子来巩固和解释这些概念,接着引入接入控制,通过分析保护信道控制过程,研究在新业务的阻塞概率和切换掉话概率之间寻求一个可能的折中。注意,在进行这种介绍性的分析时没有必要引入特定的小区形状。在后面的章节里,用不同的移动台移动模型,先是对一维小区模型,然后对二维小区模型进行等价的性能分析。

现在考虑在一个蜂窝无线网络中有一个移动终端正在呼叫。这个呼叫将在移动台当前所在的小区里发起,或者在移动台穿越小区边界的时候已经从一个相邻小区中切换过来。用参数 $1/\mu$,以秒为单位,设计一个呼叫的平均时间长度;这个时间常称为平均呼叫持续时间。假定通话时间或者说持续时间 T_n 是一个服从指数分布的随机量。很多年以来,它都是电话呼叫的常用模型。对于这种指数的情况,通话时间的概率密度函数 $f_n(t)$ 通过 $f_n(t)=\mu\exp(-\mu t)$ 确定,其中 t 用秒表示。相应的分布函数 $F_n(t)=\text{Prob}(T_n\leqslant t)$,等于 $1-\exp(-\mu t)$。

指数分布有个很有趣的特性:无记忆的性质(Papoulis,1991)[①]。它意味着一个事件的发生概率和前面的事件独立。比如说,在通话时间或持续时间服从指数分布的情况下,一个呼叫在任何给定时间上终止的概率都是一样的,也不管这个呼叫已经持续多长时间了。具有这种无记忆特性的随机变量称为马尔可夫随机变量。指数分布的随机变量是马尔可夫随机变量的一个特例。可以发现,在无记忆特性被发现是相当准确的且被引入以后,概率分析就被极大地简化了。

假定进行这种呼叫的移动台在当前小区已经分配到了一个信道,或者是切换到了本小区,或者是对一个新的呼叫请求的回应。同时,也假定移动台在当前小区内部移动(移动终端的真实定义),在进入一个相邻小区时请求分配一个新的信道,如果呼叫在进入新的小区时还未完成的话。这就是我们所说的切换过程。在切换之前,在给定小区里花费的时间 T_h 称为小区驻留时间。为了简化分析,在本节里假定这个(随机的)时间也是服从指数分布的,参数是 η,那么平均驻留时间就是 $1/\eta$。在下面的章节里,用移动性和小区几何形状的简单模型来决定驻留时间分布。知道了通话时间分布和驻留时间分布,我们就可以确定在一个给定小区里某一个信道保持或者被占用的时间长度 T_c 的分布。要想计算像呼叫阻塞概率和切换掉话概率之类的重要参数,T_c 这个量是必需的。

为了确定信道保持时间的分布,注意信道保持时间是驻留时间和通话时间的较小值;比如,如果移动台在一个给定小区内呼叫没有完成,信道保持时间就是驻留时间;如果移动台还在小区内的时候就已经完成呼叫,它就是通话时间(注意,这个判断以上面提到的无记忆假设为前提,即通话时间服从指数分布)。所以,有这样的关系

$$T_c=\min(T_h,T_n) \tag{9.1}$$

① Papoulis, A. 1991. *Probability, Random Variables, and Stochastic Processes*, 3rd edn, New York, McGraw-Hill.

现在定义这 3 个独立的随机变量的各种分布函数,分别为 $F_c(t)=\text{Prob}(T_c\leqslant t)$, $F_h(t)=\text{Prob}(T_h\leqslant t)$和 $F_n(t)=\text{Prob}(T_n\leqslant t)$。可以根据初等概率理论(Papoulis,1991)证明 3 种分布函数满足下面的关系:

$$1-F_c(t)=[1-F_h(t)]\cdot[1-F_n(t)] \tag{9.2}$$

这个方程将一个互补分布函数同其他两个互补分布函数的乘积联系起来。

正如上文假设的那样,由于呼叫持续时间和驻留时间服从指数分布,现在考虑指数分布随机变量的特殊情况。利用式(9.2),请读者证明信道保持时间也是服从指数分布的,并且它的参数 μ_c 是其他两个参数的和。这就是在这个介绍性的讨论中,驻留时间或者在小区中两次切换的时间被假定为一个指数分布变量的原因。假设平均信道保持时间为 $1/\mu_c$,则有

$$1/\mu_c=1/(\mu+\eta) \tag{9.3}$$

($1/\mu$ 和 $1/\eta$ 分别是平均通话时间和平均驻留时间)。

同样,我们也可以确定一个切换的概率 P_h。对于指数分布的随机变量来说,它被证明等于 $\eta/(\eta+\mu)$。为了展现这个简单的结论,注意这仅仅是呼叫持续期(保持时间)大于驻留时间的概率,否则就不会发生切换。先计算呼叫持续期比驻留时间的一个指定值大的概率,以那个指定值为条件,然后在驻留时间的全部取值范围内累加(积分),就得到了结果。详细的计算过程如下:

$$\begin{aligned}P_h=\text{Prob}(T_n>T_h)&=\int_0^{\infty}f_h(T_h)\left[\int_{T_h}^{\infty}f_n(t|T_h)\mathrm{d}t\right]\mathrm{d}T_h\\&=\int_0^{\infty}\eta e^{-\eta T_h}\left[\int_{T_h}^{\infty}\mu e^{-\mu t}\mathrm{d}t\right]\mathrm{d}T_h\\&=\eta/(\eta+\mu)\end{aligned} \tag{9.4}$$

式(9.4)第 3 行的结论是通过计算第 2 行的积分得到的[表达式 $f_n(t|T_h)$正是通话时间以驻留时间的给定值 T_h 为条件的密度函数]。

现在,我们来看一下利用这些结论的例子。但必须首先注意驻留时间的指数分布参数 η 也可以看成是穿越小区的平均频率,平均驻留时间是 $1/\eta$。这个频率就是在第 8 章的 8.4 节中讨论位置区域的时候曾经提到的那个量:一个移动台以 V 为速率穿越一系列面积为 S,周长为 L 的小区时的平均频率。在那里,曾经指出这个现在称为 η 的量由 Jabbari(1996)[①]给出

$$\eta=VL/\pi S \tag{9.5}$$

特别地,对于以 r 为半径的圆形小区,有

$$\eta=2V/\pi r \tag{9.5a}$$

假定平均的呼叫保持时间是 $1/\mu=200$ s。我们来看 3 种对应于不同小区尺寸的情况:一个半径为 10 km 的宏小区,一个半径为 1 km 的较小的小区和半径为 100 m 的宏小区。在第 1 种情

① Jabbari, B. 1996. "Teletraffic aspects of evolving and next-generation wireless communication networks," *IEEE Personal Communications*, 3, 6(December), 4-9.

况下,假定车辆以平均时速为 60 km 的速度运动;在另外两种情况下,假定移动的小区终端被平均时速为 5 km 的步行者持有。读者也可以考虑一下其他的情况。我们将数字代入上面的表达式就会得到下面的结果:

1. 宏小区,半径 r 为 10 km,速率 V 为 60 km/h

   两次切换之间的平均时间,或者平均驻留时间为

$$1/\eta = 940\ \text{s}$$

   切换的概率为

$$\eta/(\eta+\mu) = 0.17$$

   平均信道保持或者占用时间为

$$1/\mu_c = 170\ \text{s}$$

2. 较小小区,半径 r 为 1 km,速率 V 为 5 km/h

   两次切换之间的平均时间为

$$1/\eta = 1100\ \text{s}$$

   切换的概率为

$$\eta/(\eta+\mu) = 0.15$$

   平均信道保持或者占用时间为

$$1/\mu_c = 170\ \text{s}$$

3. 宏小区,半径 r 为 100 m,速率 V 为 5 km/h

   两次切换之间的平均时间为

$$1/\eta = 110\ \text{s}$$

   切换的概率为

$$\eta/(\eta+\mu) = 0.64$$

   平均信道保持或者占用时间为

$$1/\mu_c = 72\ \text{s}$$

当然,这些结果与所期望的相吻合。第 1 种和第 2 种情况非常相似:第 2 个小区的半径是第 1 个小区的 1/10,但是移动的速度降低得更多,所以得到的几个数字几乎一样。在这两种情况下,横穿小区的时间相对较长,因此相对应的两次切换相隔的时间也较长。由于这个时间比平均通话时间要长得多,因此切换的概率相对较小(很有可能在移动台离开小区之前就已经通话完毕,因此不需要切换)。所以,信道占用时间原则上由平均通话时间来决定。现在比较一下第 2 种和第 3 种的情况。这里小区的尺度减小为原有的 1/10,但是移动终端的速度保持不变。穿越小区的时间和发生切换的时间也将以相同的比例缩短。切换的概率也就相应地提高,通话保持时间不变。在同一个小区里的信道占用时间现在缩短很多,因为切换时间已经下降得比通话保持时间短。

现在来讨论切换业务速率,以及它和新业务到达速率之间的关系。要考虑流量的连续性(Jabbari,1996),图 9.1 画出了一个移动通信系统任意的一个小区。就像在本章开头指出的那样,假定它们都是平稳并且同分布的,这就意味着各个小区具有同样的平均流量强度和密度。[实际上,在给出式(9.5)的时候就悄悄地引入了这个假设。在推出这个表达式的时候,我们假

设移动密度均匀,小区的几何形状和终端移动速度都一样]。λ_n 为平均每个小区发起的新业务的密度,以呼叫数/s 为单位。P_b 为一个新的呼叫请求被阻塞的概率,那么一个小区纯的新业务数就是 $\lambda_n(1-P_b)$。现在平均切换的密度 λ_h,也是以呼叫数/s 为单位,因为切换就是试图从一个小区的相邻小区进入该小区。平均来说,由于假设特性一样,在离开一个小区进入相邻小区的时候发生切换的个数,一定是同一个值 λ_h。由于新的小区缺少信道,一个切换请求被拒绝的概率为 P_{hf},那么接入新小区的切换数就是 $\lambda_h(1-P_{hf})$。这些数量在图 9.1 中都指示出来了。从简单的流的角度来考虑,离开小区的平均呼叫个数 λ_h 肯定等于切换的平均数,包括新的呼叫请求和进入小区时的切换。后者由进入小区和获得信道的呼叫数量乘以它们将要切换的概率 P_h 得到。这种状态就是流量平衡的一个例子。在以后讨论呼叫准入的时候,还要给出另外一个例子。在满足指数统计特性的时候,式(9.4)已经计算了切换的概率 P_h 得到。使进入小区需要切换的呼叫数量和离开小区的数量相等,就得到了下面的式子

$$\lambda_h = P_h[\lambda_n(1-P_b)+\lambda_h(1-P_{hf})] \tag{9.6}$$

用新业务率λ_n来表示切换率λ_h,就会得到

$$\lambda_h = \frac{P_h(1-P_b)}{[1-P_h(1-P_{hf})]}\lambda_n \tag{9.7}$$

如果阻塞和切换的掉话率足够小($P_b, P_{hf} \ll 1$),那么就得到了两个业务强度的近似关系

$$\lambda_h \approx \frac{P_h}{(1-P_h)}\lambda_n, \quad P_b, P_{hf} \ll 1 \tag{9.7a}$$

特别地,对于指数分布的特殊情形,有

$$\lambda_h \approx \frac{\eta}{\mu}\lambda_n, \quad P_b, P_{hf} \ll 1 \tag{9.7b}$$

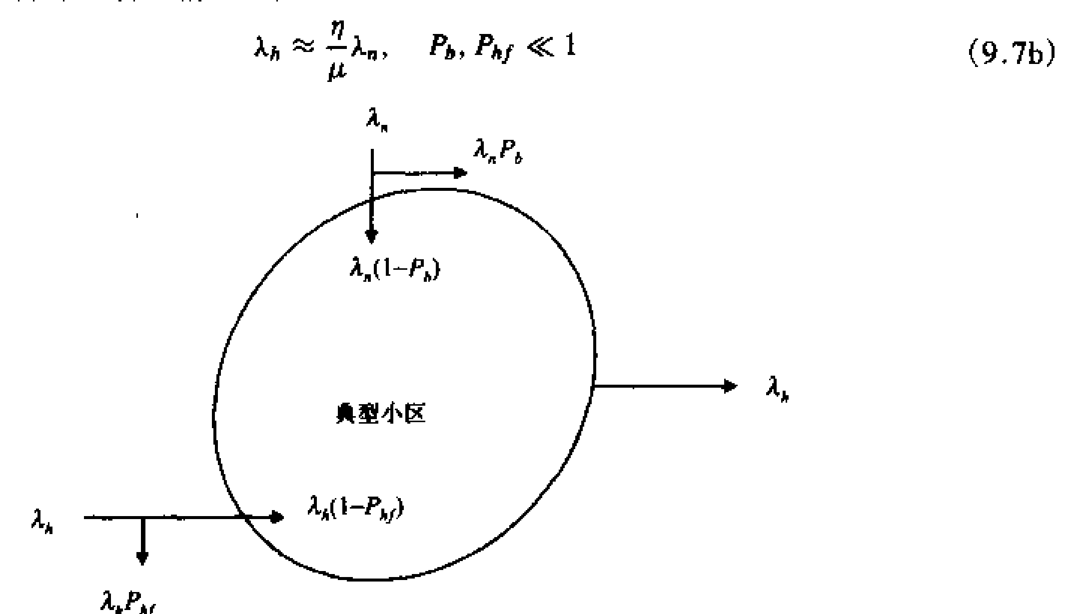


图 9.1 业务流入并流出一个小区

考虑上面所说的两个例子,分别是宏蜂窝和微蜂窝系统。把式(9.7b)代入这两种情况,在宏小区的例子中得到了 $\lambda_h = 0.21\lambda_n$,在微小区中得到了 $\lambda_h = 1.8\lambda_n$。很明显,对于所选择的数字而言,由于在微小区中切换的概率很高,新业务就会产生大量的切换。因此,在那种情况下,平均切换率几乎是新业务请求率的 2 倍。

定义 P_{hf} 为当一个终端通话时进入一个新小区而被拒绝切换的概率。在小区很小的情况下,一次通话过程中可能有很多次切换。对用户来说,一个更重要的指标是从小区到小区在一个移动终端移动的时候最终掉话的概率 P_d,它也称为"被迫终止概率"(Hong 和 Rappaport,1999)①。注意到它是一个切换业务在第1次被丢掉后,或在第2次被丢掉后,或在第3次被丢掉后,依次类推,概率的总和。第1个概率是 P_hP_{hf},切换发生的概率乘以由于容量不够被拒绝的概率。第2次切换以后掉话的概率是 $P_h^2(1-P_{hf})P_{hf}$ 等。把这些概率都加起来,就得到被迫终止概率

$$\begin{aligned} P_d &= P_hP_{hf} + P_h^2(1-P_{hf})P_{hf} + P_h^3(1-P_{hf})^2P_{hf} + \cdots \\ &= \frac{P_hP_{hf}}{[1-P_h(1-P_{hf})]} \end{aligned} \tag{9.8}$$

如果切换掉话概率 P_{hf} 很小,那么被迫停止概率就近似等于

$$P_d \approx \frac{P_h}{1-P_h}P_{hf}, \quad P_{hf} \ll 1 \tag{9.8a}$$

因此,被迫终止概率同切换掉话概率成正比,这正如所期望的那样。

现在看看呼叫阻塞概率 P_b 和切换掉话概率 P_{hf} 是由哪些因素决定的。已经说过,如果切换和新的呼叫被同等对待,那么这两个概率就是一样的,在第3章和第4章所介绍的厄兰B阻塞概率公式已经给出了这个概率(参阅式(3.7)和式(4.6))。在第3章中,我们用 N 来表示一个小区内信道的数量;而在第4章中,我们则用字母 m。这一章中,我们将用 m 来表示。然而,这里的讨论有个基本的不同。前些章节在计算业务强度时,忽略了切换,暗中关注于新业务请求,厄兰参数 A 用于计算阻塞概率。然而,如果在计算厄兰参数 A 时,在新呼叫业务强度中加入切换业务强度 λ_n,那么在那里所得到的阻塞概率 λ_h 的结果仍然是有效的。在小的小区中,由于多重切换非常普遍,所以这一点显得非常重要。已经看到,在那种情况下切换的流量已经超过了新的呼叫请求。然而,把这两个量加起来的确出现了一个问题。从式(9.7)看到,既然切换请求率和新业务的请求率成正比(毕竟,一个切换是以一个新的呼叫开始的),那么假设切换与新的呼叫独立对不对呢?答案是肯定的。因为在计算平均流量的时候,已经悄悄地在很多独立产生的呼叫上取了平均。一个正在进行通话的终端到达一个新的小区,在那个小区里请求分配一个信道(切换),很明显,和那个小区中正在通话的终端是无连接的。一个给定小区的新的呼叫是独立于到达那个小区的切换的。平均说来,切换的到达率和新业务的到达率成正比,但是在确定用于计算厄兰B阻塞概率的小区总流量的时候,两个比率可以加在一起。

如果我们通过区分两种类型的流量来提供切换业务的优先级,正如现在将要做的那样,那么在处理切换业务的时候就好像独立于新的流量,然后在计算阻塞概率和切换掉话概率的时候分别考虑两种平均流量。然而,平均的切换率仍旧和新业务的流量成正比,用式(9.7)可以计算。在本节里,只讨论提供优先级给切换业务的保护信道方法(Hong 和 Rappaport,1999)。

---

① Hong, D. and S. S. Rappaport. 1999. "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoff procedures," *IEEE Transactions on Vehicular Technology*, VT-35, 3 (August), 77-92. See also CEAS Technical Report No. 773, 1 June 1999, College of Engineering and Applied Sciences, State University of New York, Stony Brook, NY 11794.

在下一节，再介绍一种方法以进行比较。保护信道的方案就是在一个小区全部空闲的信道中预留一定数量的信道只给切换使用。在给定的小区里，有 m 个信道。其中，假设 g 只能用于切换。这就是说，全部 m 个信道都可用于切换，只有 $m-g$ 个信道可用于新的呼叫。假定新业务和切换业务的到达服从泊松分布。泊松分布和指数分布密切相关：如果到达服从泊松分布，那么到达之间的（随机）时间就服从指数分布（Papoulis，1991；Schwartz，1987）[①]。泊松分布是马尔可夫无记忆特性的另一个例子。

泊松分布的呼叫到达假设用于研究电话流量已经很多年了，在那个环境中已被很好地证实。泊松分布统计特性基于下面的假定：

考虑一个非常小的时间间隔 Δt，从 t 到 $t+\Delta t$[见图 9.2(a)]。

1. 到达概率（事件发生）是 $\lambda\Delta t \ll 1$；也就是，它很小，并且与时间间隔成正比；
2. 在相同的间隔没有到达的概率是 $1-\lambda\Delta t$，也就是，在那个间隔里，多于一个的到达的概率是可以忽略的；
3. 到达是无记忆的：在时间 t 之前到达的数量对 t 到 $t+\Delta t$ 之间到达的概率没有影响。在那个间隔里发生的事情对那个时间以后的到达没有影响。

可以证明到达时间的间隔 τ[见图 9.2(b)]是一个具有指数概率密度为 $\lambda e^{-\lambda\tau}$ 的随机变量，平均时间间隔是 $1/\lambda$（Papoulis，1991；Schwartz，1987）。也可以证明在一个更长的间隔 T 中到达的数量服从泊松分布[见图 9.2(b)]

$$\text{Prob}(\text{在 } T \text{ s 间隔内，到达的数量 } n) = (\lambda T)^n e^{-\lambda T}/n!$$

（在 T s 内没有到达的概率是 $e^{-\lambda T}$；一个到达的概率是 $\lambda T e^{-\lambda T}$，等等）很容易证明，在间隔 T 内到达的平均数量是 λT，平均到达率用单位时间内到达的数量来表示，正好是比例系数 λ。因此在假定 1 和 2 中，定义泊松分布统计特性的时候使用的比例因子 λ 被证明是到达的平均率；它的倒数 $1/\lambda$ 代表到达之间的平均时间。

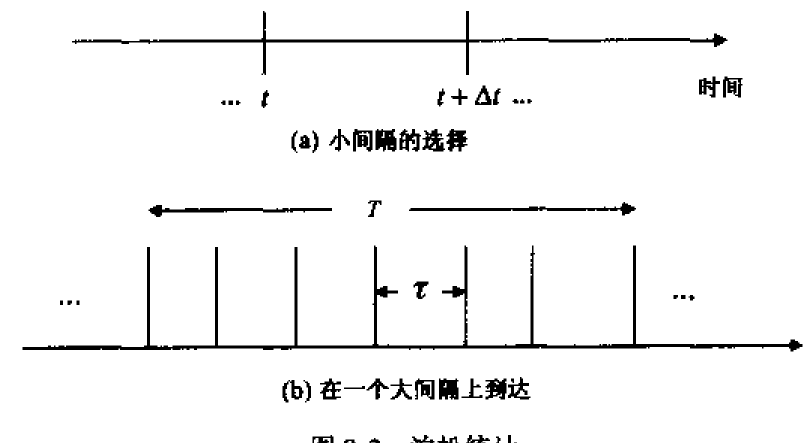


图 9.2　泊松统计

对于新的呼叫和切换到达，在用保护信道的方法计算切换掉话率和阻塞概率的时候，使用泊松分布假设。这种方法定义了一个新业务的准入过程：当还有空闲的信道为它们服务的时候，新的呼叫就被接入。根据定义，这必须是 $m-g$ 个信道中的一个，这些信道可用于新业务

---

① Schwartz, M. 1987. *Telecommunication Networks: Protocols, Modeling, and Analysis*, Reading, MA, Addison-Wesley.

请求和切换。正如上面所提到的，g 个保护信道只能用于切换请求。这称为一个准入控制过程，因为通过控制新业务的进入——没有空闲信道就阻塞它们，就可以更好地服务现有的通话和切换业务的请求。很明显，随着 g 的增加，切换掉话的概率就会降低。同时，新业务的阻塞概率就会上升。我们将看到，对于一个给定的业务强度，切换掉话概率的降低一般将会比新业务阻塞概率的升高更明显，这是一个非常好的特性。

在分析马尔可夫(无记忆)系统的时候，通常采用一种本地流量平衡的方法，这种方法也可以用来分析保护信道的准入控制过程。注意，小区中有 m 个空闲信道，系统在任何时候都处于 $m+1$ 种状态中的一种，每种状态对应一个特定的信道占用情况，包括零占用的情况。由于对呼叫到达和呼叫保持时间做了无记忆的泊松/指数假设，可以计算 $m+1$ 个状态每个状态的稳态概率。在计算中很容易找出阻塞和切换掉话状态的占用概率。在第3章推导厄兰B阻塞概率公式(3.7)，第4章研究动态信道分配性能的时候用到的阻塞状态概率，在推导它们的时候都用到了同样的方法。一个有辨识能力的读者会发现，我们在那些章节中讨论材料的时候显式地引入了泊松分布统计特性假设。在那里也显式地提到了马尔可夫链的概念。实际上，这里讨论的结果在 $g=0$ 也就是没有保护信道的时候就会退化为厄兰B公式，在前面的章节里没有加以证明。推迟了第3章的细节分析是为了简化那里的讨论。因此，对于那些不熟悉电话流量理论的人来说，本章的分析提供了在第3章中缺失的部分[很有趣的是保护信道的概念和它的分析与前面的中继线预留过程很类似，中继线预留在有线电话网中被用来处理交替路由流量(Schwartz，1987)。在电话系统行话中，一个干线就是一个信道，它处理网络中惟一连接的电话呼叫。在现代的数字电话系统中，一个干线对应于一个重复帧结构的一个时隙，同TDM无线系统中的完全一样，比如本书前面描述的IS-136和GSM)。

图9.3画出了保护信道进程的状态图，如果当前已经有至少 $m-g$ 个呼叫，新的呼叫就会被阻塞。因此新业务的阻塞状态就是 $(m-g),(m-g+1),\cdots,m$。如果 m 个信道都被占用了，切换业务就会阻塞掉。现在定义 $p_j, j=0,1,2,\cdots,m$ 为状态 j 的稳定占用概率，要计算 $m+1$ 个状态概率。一旦知道了这些，就可以通过在新业务的阻塞状态上累加来得到新业务的阻塞概率。切换掉话概率是 p_m。该图表也表明了状态之间各种不同的转换概率。随着分析的深入，为每个状态建立流量平衡方程，将会对这些转换概率了解得更多。但是考虑在图9.3中指示的这些转换概率，上(右)下(左)各个方向。注意到从0到 $(m-g-1)$ 状态向上的概率是 $(\lambda_n+\lambda_h)$。在这些状态上，新的呼叫和切换业务都可以获得信道。这个转换概率起因于两种类型的呼叫都有到达的概率，从而使状态向右移一个(在同样小的时间间隔内同时到达两个，这种可能性已经被排除了，因为泊松分布的假设。在采纳的模型中，只有相邻状态的转换才是可能的)。然而，一旦到达了 $(m-g)$ 状态，只有切换才能获得剩余的信道。因此，只有在 $(m-g)$ 到 $(m-1)$ 状态中，右移一个状态的概率是 λ_h。现在考虑左移方向，一个占有信道的通话结束了通话，或者切换到了其他小区，总之让出了那个信道，系统就左移一个状态。实现这个的概率就是参数 μ_c，它是先前讨论的平均信道保持时间的倒数。但是，随着正在通话的呼叫数量的增加，也就是说，向状态数字大的方向移动，通话让出信道的概率也相应增大：如果有 j 个通话正在进行，在一个很小的时间间隔中，这些通话让出信道的概率很有可能是只有一个通话时的 j 倍。然而，由于泊松假设的前提，在任意小的时间间隔内，只有一个通话可以结束或者切换。假定系统处于状态 j，那么让出信道的平均率就是 $j\mu_c$，在图中也是这样表示的。

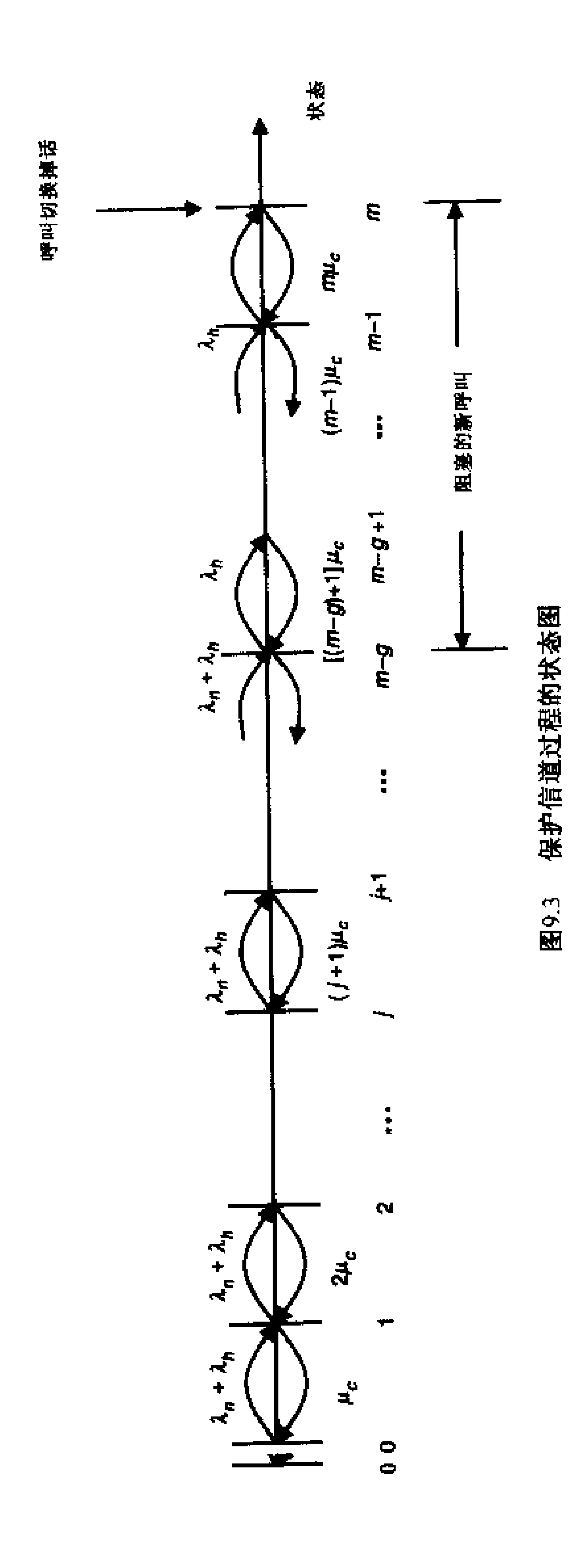


图9.3　保护信道过程的状态图

我们用这个状态转移图来确定$(m+1)$个状态的概率 $p_j, j=0,\cdots,m$。前面提到过，通过引入本地平衡方程，一个状态一个，这些概率很容易计算。比如，考虑一下状态 0，没有信道占用的状态。系统跳离这个状态的惟一办法就是一个新业务请求或者切换请求到来，然后系统就会跳到状态 1，只有一个信道被占用的状态。再次注意，正如上面所说，由于引入泊松假设，在任意一个小的时间间隔内，最多有一个任意类型的呼叫到来。那么，离开这个状态的平均率就是$(\lambda_n+\lambda_h)$，呼叫到来的概率之和乘以处于这个状态的概率 p_0。现在考虑进入这个状态。只有一种可能就是系统只有一个通话在进行，且那个通话完成了或者切换了。这时进入状态 0 的平均率就是呼叫完成或者切换的平均率 μ_c 乘以系统处于状态 1 的概率 p_1。处于统计平衡的系统离开和进入一个状态的平均率应该是相等的。使它们相等，就得到

$$(\lambda_n+\lambda_h)p_0=\mu_c p_1 \tag{9.9}$$

这个表达式是第 1 个流量平衡方程。把 p_1 表示出来，有

$$p_1=\frac{(\lambda_n+\lambda_h)}{\mu_c}p_0 \tag{9.10}$$

现在考虑状态 1。这个状态有两种独立的进入方法：从状态 2，两个通话中的一个让出信道，或者从状态 0，任意一种类型的呼叫到达。参照图 9.3，把进入状态 1 的两种可能性加起来，就得到了进入状态 1 的平均率 $2\mu_c p_2+(\lambda_n+\lambda_h)p_0$。用同样的方法，离开状态 1 的平均率，或者一个通话结束或切换而跳到状态 0，或者任意类型的一个呼叫接入而跳到状态 2，很显然就是$(\lambda_n+\lambda_h+\mu_c)p_1$。为了使到达和离开状态 1 平衡，令这两个转换率相等，有

$$(\lambda_n+\lambda_h+\mu_c)p_1=(\lambda_n+\lambda_h)p_0+2\mu_c p_2 \tag{9.11}$$

利用式(9.9)在方程的两边各消去一项，很容易简化这个平衡方程。然后，将剩余的方程用 p_1 表示 p_2，但是可以消除 p_1，利用式(9.10)直接把 p_2 和 p_0 联系起来。这样得到

$$p_2=\frac{(\lambda_n+\lambda_h)^2}{2\mu_c\cdot\mu_c}p_0 \tag{9.12}$$

现在通过定义业务强度因子 $\rho\equiv(\lambda_n+\lambda_h)/\mu_c$ 简化表示法。这个因子和厄兰业务强度因子 $A=\lambda/\mu$ 相似，即为在第 3 章中定义的平均流量到达率和平均呼叫保持时间的乘积。得到

$$p_2=\frac{\rho^2}{2}p_0 \tag{9.12a}$$

用这种方法对每一个状态 $j\leqslant(m-g-1)$建立一个平衡方程，使每个状态的离开和到达相等，以保持统计的均衡。留给读者证明下面的结果，对于状态 $j, 1\leqslant j\leqslant(m-g)$，状态概率为

$$\begin{aligned}p_j&=\frac{\rho^j}{j!}p_0,\quad 1\leqslant j\leqslant(m-g)\\ \rho&\equiv\frac{(\lambda_n+\lambda_h)}{\mu_c}\end{aligned} \tag{9.13}$$

现在考虑 j 从$(m-g)$到 m 的几个状态。再次对每个状态写出平衡方程，然而，要注意现在处于保护信道领域，因此只有切换业务才能接入信道。所以，在每个状态，流量到达率是 λ_h，如图 9.3 所示。与前面的非保护信道状态相比，这是状态图中惟一的改变。再次建立平衡方程，每个状态一个，从$(m-g)$开始。定义流量参数 $\rho_h\equiv\lambda_h/\mu_c$，读者可以证明这些方程可以被写成下面的形式：

$$p_{m-g+k} = \frac{\rho_h^k}{(m-g+k)(m-g+k-1)\dots(m-g+1)} p_{m-g},$$
$$\rho_h \equiv \frac{\lambda_h}{\mu_c}, \quad 1 \leqslant k \leqslant g \tag{9.14}$$

利用式(9.13),可以将式(9.14)出现的状态概率 p_{m-g} 用 p_0 表示出来。同时用 j 代替 $m-g+k$,就可以简化为

$$p_j = \frac{\rho_h^{j-(m-g)} \rho^{m-g}}{j!} p_0, \quad m-g+1 \leqslant j \leqslant m \tag{9.14a}$$

这样利用平衡方程就完成了 m 个状态概率 $p_j, 1 \leqslant j \leqslant m$ 的计算。所有的状态概率都用 p_0 表示,保护信道系统有 m 个空闲信道的概率。现在还要确定这个概率。由于 $m+1$ 个状态概率之和为 1,所以这个概率的计算很简单。把式(9.13)和式(9.14a)在各自的变化范围内相加,就得到在式(9.13)和式(9.14a)中用到的 p_0 的值

$$\frac{1}{p_0} = \sum_{j=0}^{m-g} \frac{\rho^j}{j!} + \rho^{m-g} \sum_{j=m-g+1}^{m} \frac{\rho_h^{j-(m-g)}}{j!} \tag{9.15}$$

现在很容易计算新业务的阻塞概率和切换掉话概率,就像先前指出的那样,使用近似的状态概率的形式。特别地,立即得到了切换掉话概率 P_{hf} 为

$$P_{hf} = p_m = \rho^{m-g} \frac{\rho_h^g}{m!} p_0 \tag{9.16}$$

p_0 在式(9.15)中给出。

新业务的阻塞概率 P_b 为

$$P_b = \sum_{j=m-g}^{m} p_j = \rho^{m-g} \sum_{j=m-g}^{m} \frac{\rho_h^{j-(m-g)}}{j!} p_0 \tag{9.17}$$

在运用这些方程进行数值计算的时候有一个问题。业务强度参数 ρ 和 ρ_h 都是以切换业务到达率 λ_h 的形式定义的。从式(9.7)看出,这个量取决于 P_{hf} 和 P_b 两个量! 式(9.7a)指出,只有当这两个概率值很小的时候,λ_h 才和它们独立。答案是进行一个迭代的分析,在式(9.16)和式(9.17)出现 λ_h 的表达式中,先设定 P_{hf} 和 P_b 都等于 0,得到概率的值,再计算 λ_h,在式(9.16)和式(9.17)中用这个新值,一直重复到方程收敛。马上就要指出这种分析的结果,但先要回到厄兰 B 方程上来。令保护信道 $g=0$,从式(9.16)和式(9.17)中很容易看出,阻塞和切换掉话概率收敛于相同的值$(\rho^m/m!)p_0$。p_0 的表达式(9.15)也进行了简化。读者可以证明阻塞概率实际上由厄兰 B 分布给出

$$P_b\Big|_{g=0} = \frac{\rho^m/m!}{\sum\limits_{j=0}^{m} \rho^j/j!} \tag{9.18}$$

注意,这正是第 3 章式(3.7)的形式,如前所述。由于业务强度的确依赖于 λ_h,它又反过来依赖 P_b,所以在进行包括切换的计算时仍然必须重复这个方程。

考虑一下先前描述的保护信道准入控制方案对宏小区和微小区的应用例子。回忆一下,在那两个例子中切换的概率 P_h 分别是 0.17 和 0.64。表 9.1 给出了计算这两种情况下新业务

先考虑宏小区的例子。很明显，在这里掉话率上的改善比阻塞率的增加要显著得多。注意 $\rho_n=5$，当没有使用保护信道的时候，两种概率都是0.042。指定10个信道中的1个为保护信道，切换掉话率就会降低到0.0068，不到原来的1/6。阻塞概率升高到0.079，不到原来的2倍。如果指定两个保护信道的话，切换掉话概率将进一步降低到0.001，为原来的1/7，而阻塞概率升高到0.13，增大程度小得多。当然，还可以进一步增加保护信道的数量，但那样就会导致阻塞概率非常高。注意，特别地，$g=3$ 时，切换掉话概率降到一个非常小的数值，当业务强度在5到8之间时，该值徘徊在 10^{-4} 和 5×10^{-4} 之间。然而，相对应的新业务的阻塞概率徘徊在0.19到0.39之间。当然，增加信道的数量会降低两个概率，因此，保护信道的引入会带来希望的结果：正在进行的通话在切换时掉话的概率更低。所导致的阻塞概率的升高，如果不太高的话，可以被认为是对进行中通话服务质量改善的一个可以接受的代价。

对于这种每个小区有10个信道的宏小区来说，怎样才是一个好的设计呢？在一个新呼叫业务强度 $\rho_n=5$ 上进行操作可能会提供一个可以忍受的性能，也就是说，一个相对好的服务等级。如果把一个信道拨给切换专用的话，切换掉话率大约是0.7%，而相应的新业务阻塞概率大约是8%。注意，这里的业务强度数字不是3.3节定义的以厄兰为单位的业务强度。那里的厄兰定义是 $A\equiv\lambda_n/\mu$，$1/\mu$ 是呼叫保持时间。这里用到了参数 $\rho_n\equiv\lambda_n/\mu_c$，$1/\mu_c$ 是平均信道保持时间，和 $1/(\mu+\eta)$ 相等，$1/\eta$ 是平均小区驻留时间。厄兰负载实际上比这里得到的业务强度参数要高。特别地，有 $A=\rho_n(1+\eta/\mu)$。对于这里的宏小区的例子有，$1/\mu=200$ s，$1/\eta=940$ s，$A=1.21\rho_n$。当 $\rho_n=5$ 时，每个小区允许的新业务数 $A=6$ 厄兰（当然，本书中讨论的3个第二代系统都允许比这高得多的业务强度，因为它们的容量要高得多，请看第6章）。通话时间是200 s，6-厄兰转换成一个平均呼叫到达率——每200 s 6个呼叫或者每分钟1.8个呼叫。把保护信道的数量增加到2个或3个，会导致一个非常大的阻塞概率，正如上面所说。增加新呼叫业务强度到 $\rho_n=6$ 或者更高，也会导致更高的阻塞概率。

现在讨论第2种微小区的例子。表9.1表明，所允许的新业务的业务强度非常小，因为切换增加很多。特别地，在新呼叫业务强度为 $\rho_n=2$ 时，如果没有保护信道的话，阻塞概率和切换掉话概率都是2.5%。这个值对于新业务提供了很好的阻塞性能，但是对于正在进行中的通话来说，掉话概率太高了。如果在10个信道中指定一个作为切换专用的话，切换掉话率就降低到1.5%，而阻塞概率却升高到6%，这是一个可以接受的值(在这种情况下，增加信道的数量会更有帮助！所以值得重复一下，$m=10$ 这个选择是任意的，为了提供一个简单的容易计算的例子，也是为相关性能的讨论提供一个工具)。使用两个保护信道会将掉话概率降到0.8%，但阻塞概率会增到10%。如果新呼叫业务强度增大到 $\rho_n=3$，且没有指定保护信道的话，阻塞和切换掉话概率都是8.7%。若指定一个保护信道供切换专用，则切换掉话率就会降到4.9%，而阻塞率会增到一个相当高的数值17%。所以，如果只有10个可用的信道的话，那么最好保持在最高的业务强度 $\rho_n=2$，在10个中选定1个作为切换专用信道。在上面厄兰流量的讨论中，以厄兰为单位的流量是 $A=5.6$ 厄兰(回想一下，对于微小区模型，$\eta/\mu=1.8$)。这个数字和宏小区模型并没有很大的差别。尽管由于小区尺度的减小，切换业务的增多，从而导致流量增大，但当呼叫切换的时候，信道保持时间较短。这两种影响可互相抵消。

在下两节中，我们通过特定的几何形状和移动模型来测试一下系统性能。从下一节开始，我们开始讨论简单一点的一维小区模型，而随后一节我们讨论二维的情况。

9.2 一维小区

在本书的前面使用过一维小区模型,先是在3.2节中计算SIR,然后是4.2节中提供功率控制算法的例子。一维的情况提供了进行计算的简单模型,同时也提供了高速公路旁边的小区的模型。本节的目标就是像前面章节那样提供性能分析,但是专门加入了移动性和一维的几何形状。由于专注于一种特定的几何形状,在分析中会有一些改变:通过利用移动速度模型结合几何形状,可以从基本原理上确定移动驻留时间的分布,而不是先前假定它是指数分布;另外,因为驻留时间的分布不再是无记忆的,对于新到达呼叫的驻留时间和切换业务是不一样的,这同样也会影响分析。这些一般的考虑与实际的几何形状相独立,因此在下一节分析二维小区的形状时也会用到。

先考虑两种独立情况下的驻留时间:新发起的呼叫和从先前的小区切换过来的呼叫。再次注意驻留时间对应于一个移动台在切换到其他小区前在本小区的时间。特别地,专注于图9.4画的一维小区的形状。小区的长是L米。再次假定同分布的情况:所有的小区都是类似的,都有同样的均匀分布的移动台数目。在小区里,所有移动台的速度也服从同样的分布(关注两种情形:所有的移动台都以同样的恒定速率移动,或者速率服从均匀分布,有个最大值)。新业务假定在小区的任何地方都是随机、均匀产生的。用符号T_N来表示一个新产生的呼叫的小区驻留时间;T_H表示切换移动台的驻留时间。相应的概率密度和分布函数分别定义为$f_N(t)$和$F_N(t)$;$f_H(t)$和$F_H(t)$。在前面的部分,假定驻留时间服从指数分布,在这两种随机变量之间没有做区分。

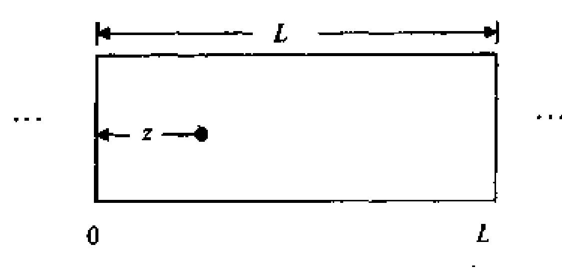


图9.4 典型小区、一维系统

用这两种密度函数计算两种驻留时间下的切换概率。回想一下前面的章节,切换的概率就是指通话时间T_n大于驻留时间的概率。用式(9.4)可以计算这两种情况,而没有假定驻留时间服从指数分布。因此,一旦得到了两种驻留时间变量的密度函数,那么各自的切换概率$P_N = \text{Prob}[T_n > T_N]$和$P_H = \text{Prob}[T_n > T_H]$,由下面的方程得到:

$$P_N = \int_0^\infty e^{-\mu t} f_N(t)\mathrm{d}t \tag{9.19}$$

和

$$P_H = \int_0^\infty e^{-\mu t} f_H(t)\mathrm{d}t \tag{9.20}$$

再次注意P_N是一个新产生的呼叫切换的概率;P_H是一个从别的小区切换过来以后再次发生切换的概率。这些概率是P_k的一般化,后者是假定驻留时间服从指数分布的切换概率,如

式(9.4)所示[式(9.19)和式(9.20)分别是各自函数的拉普拉斯变换。它们也是特征函数，因此各自随机变量的各种时刻 T_N 和 T_H，可以通过微分得到，如果它们存在的话]。

对于特定的小区形状和移动速度分布，不只是这两个概率不同，切换业务率 λ_h 和新业务的发起率 λ_n 的关系也不一样。这反过来影响了信道保持时间的分布，以后我们就能看到。但是先考虑切换业务率和新业务的发起率的关系。我们将通过扩展基于图 9.1 的如式(9.6)所示的流量平衡关系来得到这里定义的两种切换概率的关系，根据处理的是新业务还是切换业务。特别地，从一个给定小区切换出来的平均率必须和总的切换业务率相等，包括从这个小区产生的新业务和从相邻小区切换过来的呼叫，得到下面的平衡方程：

$$\lambda_h = P_H\lambda_h(1 - P_{hf}) + P_N\lambda_n(1 - P_b) \tag{9.21}$$

和前面一样，P_{hf} 是切换掉话概率，P_b 是新业务的阻塞概率，这个值取决于所选择的准入控制过程。整理式(9.21)，作为式(9.7)更一般的表达式

$$\lambda_h = \frac{P_N(1 - P_b)}{1 - P_H(1 - P_{hf})}\lambda_n \tag{9.22}$$

现在找到两个驻留时间的分布和相应的切换概率 P_N 和 P_H。

先考虑这种情况：在一个典型的一维小区里产生了新的呼叫，正如图 9.4 所表示的那样。假定当呼叫发起的时候，发起这些呼叫的移动台等可能地分布在长度为 L m 的小区的任何地方。如上所述，考虑两种移动台的速率分布。第 1 种情况简单地假定所有的移动台都以恒定的速率 V_c 移动，并且朝任何方向的可能性是相等的。第 2 种情况假定移动速率，尽管在小区内移动的时候是保持恒定的，从 0 到一个最大值 V_m 之间是随机分布的。平均来说，这些移动台的平均速率是 $V_m/2$ m/s。对每一种情况，分别计算 $f_N(t)$、$F_N(t)$ 和 P_N。特别地，对于恒定的、确定的速率，读者可以很简单地证明，小区的驻留时间密度函数是

$$f_N(t) = \frac{V_c}{L} \quad 0 \leqslant t \leqslant \frac{L}{V_c} \tag{9.23}$$

相应的分布函数是

$$\begin{aligned} F_N(t) &= \frac{V_c}{L}t \quad 0 \leqslant t \leqslant \frac{L}{V_c} \\ &= 1 \qquad t \geqslant \frac{L}{V_c} \end{aligned} \tag{9.24}$$

由式(9.19)可知，相应的切换概率是

$$\begin{aligned} P_N &= \int_0^\infty e^{-\mu t} f_N(t)\mathrm{d}t = \frac{V_c}{L}\int_0^{L/V_c} e^{-\mu t}\mathrm{d}t \\ &= \frac{V_c}{\mu L}(1 - e^{-\mu L/V_c}) \end{aligned} \tag{9.25}$$

因此，这个概率取决于标量参数 $\mu L/V_c$。随着小区宽度 L 的减小，速率和呼叫保持时间不变，一个新的呼叫切换的概率就会增大。如果速率 V_c 和平均呼叫保持 $1/\mu$ 增加，同样的观测也是正确的。随着 $\mu L/V_c$ 的减小，新业务切换的概率接近于 1。这些都与直觉相吻合。随着这个参数的增加，切换的概率迅速地接近 $1/(\mu L/V_c)$，然后随着参数的继续增加趋向于 0。

现在考虑第2种情况,移动台的速率 V 服从0到 V_m 的均匀分布。随机驻留时间 T_N 是 z/V,距离 z 是移动台在发起呼叫请求的地方和小区的边缘准备进行切换的地方两点之间的距离。显然,z 的取值在0到 L 之间。图9.4展示了一个例子,一个移动台从 z 点向左移动。然而,移动台移动的方向对于计算而言是无关紧要的。在这种情况下,随机驻留时间就是两个随机变量的比值。定义为两个随机变量比值的随机变量概率密度函数是很容易得到的,参看概率理论的基本原理一书(Papoulis,1991)。特别地,驻留时间密度函数 $f_N(t)$ 可以写为

$$f_N(t) = \int V f_z(tV) f_V(V) \mathrm{d}V \tag{9.26}$$

在这种情况下,两个随机变量 z 和 V 都是均匀分布的,$f_z(z)=1/L, 0\leqslant z\leqslant L$;$f_V(V)=1/V_m$,$0\leqslant V\leqslant V_m$,因此,$f_z(Vt)=1/L, 0\leqslant Vt\leqslant L$。代入式(9.26),得

$$\begin{aligned} f_N(t) &= \int_0^{V_m} \frac{V}{LV_m}\mathrm{d}V = \frac{V_m}{2L}, \qquad t\leqslant \frac{L}{V_m} \\ &= \int_0^{L/t} \frac{V}{LV_m}\mathrm{d}V = \frac{L}{2V_m t^2}, \quad t\geqslant \frac{L}{V_m} \end{aligned} \tag{9.27}$$

图9.5中画出了驻留时间密度函数。对密度函数积分得到相应的概率分布函数

$$\begin{aligned} F_N(t) &= \int_0^t f_N(t)\mathrm{d}t = \frac{V_m t}{2L}, \qquad t\leqslant \frac{L}{V_m} \\ &= 1 - \frac{L}{2V_m t}, \quad t\geqslant \frac{L}{V_m} \end{aligned} \tag{9.28}$$

注意,驻留时间 $T_N\leqslant L/V_m$ 的概率是0.5。对于固定速率的情况和随机速率的情况,其分布函数分别由式(9.24)和式(9.28)给出,在图9.6中已画出。固定速率 V_c 在随机速率的情况下被选择为平均速率 $V_m/2$。注意结果的相似之处。

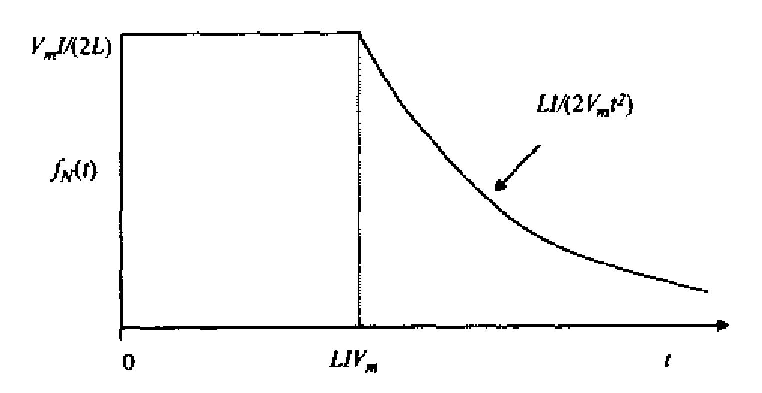


图9.5 一维系统,随机速率下驻留时间密度函数

利用式(9.19),可以计算在移动速率均匀分布的情况下,一个新的呼叫切换的概率 P_N。由式(9.19)和式(9.27)有

$$
\begin{aligned}
P_N &= \int_0^{\infty} \mathrm{e}^{-\mu t} f_N(t)\mathrm{d}t \\
&= \int_0^{L/V_m} \frac{V_m}{2L}\mathrm{e}^{-\mu t}\mathrm{d}t + \int_{L/V_m}^{\infty} \frac{L}{2V_m t^2}\mathrm{e}^{-\mu t}\mathrm{d}t \\
&= \frac{1}{2a}[1-\mathrm{e}^{-a}] + \frac{a}{2}\int_a^{\infty}\frac{\mathrm{e}^{-x}}{x^2}\mathrm{d}x
\end{aligned}
\tag{9.29}
$$

这里的参数 $a \equiv \mu L/V_m$。注意这个标量参数和式(9.25)中出现的 $\mu L/V_c$ 非常相似。那里关于变化参数的影响的讨论在这里仍旧成立:随着参数 a 的增大,通过增加小区的长度 L,减小最大速率 V_m 或者缩短呼叫保持时间 $1/\mu$,切换的概率降低。反之,随着参数的减小,切换的概率升高。

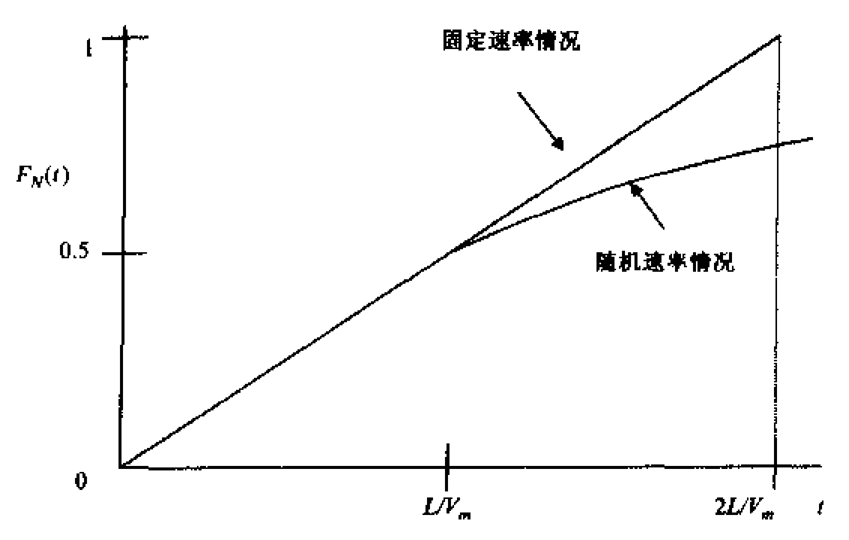


图 9.6 一维系统,$V = V_m/2$ 驻留时间概率分布函数

例子

现在提供 3 个例子,利用式(9.25)和式(9.29)来计算新业务切换的概率 P_N,通过变化小区尺寸,移动速率来使这些结论更可靠。这些例子使我们可以比较固定速率和随机速率分布的情况。在 3 种情况下,选择 $1/\mu = 200$ s。同时,选择 $V_c = V_m/2$,也就是,固定速率一直和随机速率的平均值相等。

1. 宏小区,$L = 10$ km,常数 $V_c = 60$ km/h, $V_m = 120$ km/h

$$P_N|_{\text{fixed}} = 0.316 \qquad P_N|_{\text{random}} = 0.30$$

2. 宏小区,$L = 1$ km,常数 $V_c = 60$ km/h, $V_m = 120$ km/h

$$P_N|_{\text{fixed}} = 0.86 \qquad P_N|_{\text{random}} = 0.78$$

3. 微小区,$L = 100$ m,常数 $V_c = 5$ km/h, $V_m = 10$ km/h

$$P_N|_{\text{fixed}} = 0.84 \qquad P_N|_{\text{random}} = 0.76$$

注意,速率均匀分布的时候新业务的切换概率肯定比固定速率的时候小。平均速率值的方差降低了切换的概率。第 2 种情况和第 3 种情况非常相似,因为小区的尺寸和速率几乎减小了一样的幅度。如果微小区的速率选择是 6 km/h,结果将完全一样,因为两种情况下的参数完全一样。但是,在大一点的宏小区里,也就是第 1 种情况,切换的概率会比小一些的第 2 种情况小得多。这也是理所当然的。

在小区里，驻留时间的分布和 P_N 的计算都是针对新业务到达的情况。现在重复计算，这次针对从别的小区切换过来的呼叫。回忆一下，如果驻留时间的分布是指数的或者无记忆的，那么这两种情况下的结果应该是一样的。由于这里使用了特定的小区形状，驻留时间分布不再是指数的，所以两种情况也就不一样了。参考一下一维小区的图 9.4。先考虑一下固定速率的情况。一个切换业务穿越小区要用 L/V_c s 的时间，因此驻留时间 T_H 也就是这个常数值，驻留时间的概率密度函数 $f_H(t)$ 是一个冲激函数 $\delta(t-L/V_c)$，分布函数是 $F_H(t)=0, t<L/V_c; F_H(t)=1, t\geqslant L/V_c$。切换概率 P_H 就是驻留时间 T_H 比小区保持时间 T_n 短的概率，由式(9.20)得

$$P_H = e^{-\mu L/V_c} \tag{9.30}$$

把此概率与由式(9.25)得到的新业务的切换概率相比较。一些想法表明 P_H 总是小于 P_N。这符合预期，因为一个切换业务总是需要遍历整个小区长度 L，而一个新业务平均只需要移动 $L/2$ 的距离。因此，在很大程度上，相对于一个新业务来说，一个切换过来的呼叫当到达小区的另一边需要再次切换的时候已经完成了通话。

现在考虑速率是随机的情况。在这种情况下移动的速率服从 0 到 V_m 的均匀分布。平均速率是 $V_m/2$。因为小区驻留时间总是 $T_H=L/V$，V 是一个随机变量，用常见的方法来确定 T_H 的密度函数 $f_H(t)$——通过变换随机变量(Papoulis, 1991)：

$$f_H(t) = \frac{f_V(V)}{\left|\frac{dt}{dV}\right|} = \frac{L}{V_m t^2}, \quad t \geqslant \frac{L}{V_m} \tag{9.31}$$

由式(9.20)可知，切换发生的概率 P_H 为

$$P_H = \int_{L/V_m}^{\infty} e^{-\mu t}\frac{L}{V_m t^2}dt = a\int_a^{\infty}\frac{e^{-x}}{x^2}dx \tag{9.32}$$

参数 a 定义为 $\mu L/V_m$。通过比较等价的式(9.29)，$P_N=(1-e^{-a})/2a+P_H/2$。从先前的例子中我们将要看到，$P_N>P_H$，就像上面速率固定时一样。

在表 9.2 中，比较了上面 3 种情况下的切换概率 P_N 和 P_H，对于使用的两种模型——速率恒定和均匀分布的情况，还有前面假定驻留时间服从指数分布时得到的结果。在所有的情况中，平均呼叫保持时间是 200 s。注意 $P_N>P_H$，当 $V_c=V_m/2$ 时，两种模型的几个概率相差不大。指数驻留时间时的切换概率处于速率随机模型的 P_N 和 P_H 这两个值之间。只有对于第 1 种情形，即较大的宏小区，P_H 和 P_N 才具有明显的不同。

表 9.2 切换概率

P_N		P_H		
恒定V	随机V	恒定V	随机V	$P_H\|_{指数}$
1 宏小区，$L=10$ km, $V_c=60$ km/h, $V_m=120$ km/h				
0.316	0.30	0.05	0.074	0.17
2 宏小区，$L=1$ km, $V_c=60$ km/h, $V_m=120$ km/h				
0.86	0.78	0.74	0.64	0.69
3 微小区，$L=100$ m, $V_c=5$ km/h, $V_m=10$ km/h				
0.84	0.76	0.70	0.60	0.64

P_H 和 P_N 的这些值可用于式(9.22)中,计算切换业务到达的平均率 λ_h,而 λ_h 可用新业务的到达率 λ_n 来表示,尤其当新业务的阻塞概率 P_b 和切换掉话概率 P_{hf}小到可以忽略的时候。特别地,计算比值 λ_h/λ_n,并将它们和已经得到的指数驻留时间模型的比值相比较是一件很有趣的事情。若忽略阻塞和切换掉话率,则随机速率模型和指数模型的比较结果如下:

		λ_h/λ_n	
		随机速率模型	指数驻留时间
1	10 km宏小区	0.32	0.20
2	1 km宏小区	2.2	2.1
3	100 m微小区	1.9	1.8

注意一下,尽管模型有差异,结果却是非常类似。

在做这些比较的时候,我们忽略了阻塞概率 P_b 和切换掉话概率 P_{hf}。这些重要的量是移动小区性能的实际指标,为了得到它们,必须像前面9.1节那样,定义一个准入控制过程,比如保护信道技术。实际上,不久我们将比较保护信道过程和另一个预料性的过程。但是为了分析信道保护过程,必须先确定信道占用或者保持时间,正如前面所做的那样。这里的不同就在于要考虑两种情况:对于切换业务的信道保持时间和对于新业务的信道保持时间。从这个目的出发,回到式(9.1),并将它和基于它的式(9.2)一般化。T_{HH}为切换业务的信道保持时间,T_{HN}是对应的新业务的信道保持时间。两个都是随机变量,它们的概率密度可以通过式(9.2)得到。特别地,T_{HH}一定是切换业务的驻留时间和通话时间两个中较小的那个,因为无论先出现哪种情况,小区中的信道将一直保持到移动台切换到另一个小区或者通话完成。T_n 还是表示随机的通话时间,服从指数分布 $F_n(t)$,有

$$T_{HH}=\min(T_n,T_H) \tag{9.33}$$

$F_{HH}(t)$是 T_{HH}的概率分布函数,回想一下刚完成的分析,$F_H(t)$是小区驻留时间的分布函数,它的决定因素刚才已经完成了。比较式(9.1)、式(9.2)和式(9.33),

$$1-F_{HH}(t)=[1-F_n(t)]\cdot[1-F_H(t)] \tag{9.34}$$

来源于概率论的这个方程使一个互补分布函数和另两个互补分布函数的乘积相等。

相似地,$F_{HN}(t)$是新业务的信道保持时间 T_{HN}的分布函数,有

$$T_{HN}=\min(T_n,T_N) \tag{9.35}$$

因此,

$$1-F_{HN}(t)=[1-F_n(t)]\cdot[1-F_N(t)] \tag{9.36}$$

对于任何一种情况,新业务或者切换业务,$F_c(t)$是总的信道保持时间的概率分布函数。它由两个信道保持时间的分布 $F_{HH}(t)$和 $F_{HN}(t)$加权得到,加权系数是各自产生流量的比例。一个小区实际拥有的平均新业务流量是 $\lambda_{nc}=\lambda_n(1-P_b)$;对应的小区切换流量是 $\lambda_{hc}=\lambda_h(1-P_{hf})$,$\lambda_h$ 通过式(9.22)用 λ_n 表示。总的流量就是 $\lambda_{nc}+\lambda_{hc}$,信道保持时间的分布 $F_c(t)$由加权和得到

$$F_c(t)=\frac{\lambda_{nc}}{\lambda_{nc}+\lambda_{hc}}F_{HN}(t)+\frac{\lambda_{hc}}{\lambda_{nc}+\lambda_{hc}}F_{HH}(t) \tag{9.37}$$

利用式(9.34)和式(9.36),假定通话时间的分布 $F_n(t)$是指数的,平均值是 $1/\mu$,也就是,$F_n(t)=1-\exp(-\mu t)$,该假设贯穿于本章始末,读者可以证明式(9.37)能写成下面的形式,直接包含两个小区驻留时间分布 $F_N(t)$和 $F_H(t)$:

$$F_c(t)=1-\mathrm{e}^{-\mu t}+\frac{\mathrm{e}^{-\mu t}}{1+\gamma_c}[F_N(t)+\gamma_c F_H(t)] \tag{9.37a}$$

上式中出现的参数 γ_c 定义为切换业务和新业务的流量之比 $\lambda_{hc}/\lambda_{nc}$。

在式(9.37)和式(9.37a)中,通过 λ_{hc},λ_{nc}和 γ_c 隐式出现的 P_b 和 P_{hf}取决于选定的准入控制过程。正如前面那样,一旦过程指定了以后,需要用迭代的过程来计算这些量。把这些量运用到保护信道准入过程的时候有一个问题。那里做的假设是信道保持时间服从指数分布,这样我们就得到了图 9.3 所示的状态图。基于它我们建立了本地流量平衡方程,比如式(9.11)。从式(9.37(a))容易看出,现在已经不是那个情况了。因此,采用 Hong 和 Rappaport(1999)的方法,把信道保持时间分布近似为一个等价的指数分布。选择的指数分布保留了实际的平均信道保持时间 $1/\mu_c$。平均信道保持时间就是$\int tf_c(t)\mathrm{d}t$,$f_c(t)$是信道保持时间的概率密度函数。对于表现好的函数,也可以写成$\int[1-F_c(t)]\mathrm{d}t=\int F_c^C(t)\mathrm{d}t$,$F_c^C(t)$是互补分布函数。那么,等价的指数分布就是 $1-\exp(-\mu_c t)$,平均值是 $1/\mu_c$。因此,很明显,对两个分布使用同样的平均值就等价于

$$\int_0^\infty[F_c^C(t)-\mathrm{e}^{-\mu_c t}]\mathrm{d}t=0 \tag{9.38}$$

这种近似指数的另一种解释是互补信道保持时间的分布和指数之差的积分等于 0(Hong 和 Rappaport,1999)。从式(9.38)和式(9.37a)很容易得到

$$\frac{1}{\mu_c}=\frac{1}{\mu}-\frac{1}{1+\gamma_c}\int_0^\infty \mathrm{e}^{-\mu t}[F_N(t)+\gamma_c F_H(t)]\mathrm{d}t \tag{9.39}$$

再次提示 $\gamma_c\equiv\lambda_{hc}/\lambda_{nc}$,$\lambda_{hc}=\lambda_h(1-P_{hf})$,$\lambda_{nc}=\lambda_n(1-P_b)$(Hong 和 Rappaport, 1986)。式(9.39)给出的 μ_c 值将被用于保护信道准入控制过程的性能的各种计算。

例子

一个计算 $1/\mu_c$ 的例子很有用。为了简单起见,采用固定速率模型。假定 P_b 和P_{hf}都是零。由式(9.24)可知,有 $F_N(t)=V_c t/L, 0\leqslant t\leqslant L/V_c$;$F_N(t)=1, t\geqslant L/V$,还有 $F_H(t)=0, t<L/V_c$;$F_N(t)=1, t\geqslant L/V_c$。对于 10 km 的宏小区模型,移动速率是 60 km/h,$1/\mu\approx 200$ s,由式(9.39)得 $1/\mu_c=0.75/\mu=150$ s。假定为指数切换模型,前面的部分得出了可以比较的结果是 170 s。对于 100 m 的微小区,移动速率是 5 km/h,得 $1/\mu_c=0.26/\mu=53$ s。由指数切换模型得到的值是 72 s。就像在指数切换模型里一样,两种尺寸的小区在信道保持时间上的差异是预料之中的。即使以步行的速率 6 km/h 移动,遍历小得多的微小区也要快得多。因此,平均来说,信道保持时间要短得多。

现在回到准入控制机制上来。小区为了保持较好的性能,就要引入某种准入控制过程。如果流量太高的话,在电话系统里一个常见的过程是阻塞新到达的请求。在这种情况下,流量

是用厄兰为单位，定义为呼叫到达率和平均呼叫保持时间的乘积。在前面几章里，引入阻塞概率作为一个性能指标。移动系统也引入了切换掉话率。在前面的部分提到，保持一个很低的切换掉话率是很合理的，因为一旦一个通话已经建立，而在进入一个新的小区时又被切断的话，用户会感到非常讨厌。对于一个给定的小区信道容量，要给予切换业务一定的优先级。前面部分所描述的保护信道过程正是这样做的。还有一个切换优先级过程被提出和被很多人研究——将切换业务排队，直到有空闲信道为止（Hong 和 Rappaport，1999）。还有很多准入控制过程被提出和加以比较。在这里概述一个，只针对一维的小区系统，把那种情况下的性能和保护信道方案加以比较（Naghshineh 和 Schwartz，1996）[1]。这个过程是分布式的，包括相邻小区的信息交换，对小区可以接受的用户门限进行设定，这样，对于不久就要切换的通话就可以保证得到信道。

在一维系统里考虑 3 个相邻的小区。在小区 C_n 里有个终端发起准入请求。C_l 和 C_r 是它的相邻小区，分别在 C_n 的左边和右边。P_i 是小区里有 i 个通话的概率，比如一个小区可以有 N 个通话。定义过载概率是小区里呼叫的数量超过 N 的概率，这个概率被指定为一个可以接受的值 q。在允许一个新的呼叫进入 C_n 的时候要保证这个目标。为了达到这个目标，小区的准入控制器用到了在小区 C_n，C_r 和 C_l 中当前呼叫的数量和估计将来 T s 内呼叫的数量，T s 大致等于切换的时间。所以，准入控制过程的关键是预测，估计稍后是否有空闲信道供切换。具体来说，必须满足 3 个准入条件：

1. 将来 T s 内，由于和相邻小区相互切换，小区 C_n 的过载概率必须小于 q。
2. 将来 T s 内，由于和相邻小区相互切换，小区 C_r 的过载概率必须小于 q。
3. 将来 T s 内，由于和相邻小区相互切换，小区 C_l 的过载概率必须小于 q。

在下面同分布的情况下，这些准入条件可以被量化。

p_m 是 T s 内一个小区中切换的概率。如果假定切换服从指数统计特性，平均律是 $\dot{\eta}$，则 $p_m = 1 - e^{-\eta T}$。由于同分布，切换到左右小区的概率各为 $p_m/2$。T s 内没有切换的概率是 $1 - p_m = e^{-\eta T}$，定义为 p_s。现在考虑一个给定的小区。假定当前小区中有 k 个通话，T s 以后这些通话中有 i 个没有切换的概率是二项分布

$$B(i, k, p_s) = \frac{k!}{i!(k-i)!} p_s^i (1 - p_s)^{k-i} \tag{9.40}$$

这个分布的均值是 kp_s，方差是 $kp_s(1 - p_s)$。当 k 很大的时候，它可以近似为一个具有同样均值和方差的高斯分布（二项分布就是二元随机变量的和。这里的二元条件就是一个通话是否发生切换。根据概率的中心极限定理，一系列随机变量和的分布趋向于高斯或者正态分布）。

现在考虑小区 C_n，当新业务来临时，准入控制将被激活。假设在新的呼叫到来时，包括新的呼叫一共有 n 个呼叫正在通话。假定这一时刻 C_r 小区和 C_l 小区各有 r 和 l 个呼叫正在进行。基于上面的讨论，读者可以证明，T s 后小区 C_n 中呼叫的数量大致服从一个均值为

---

[1] Naghshineh, M. and M. Schwartz. 1996. "Distributed call admission control in mobile/wireless networks," *IEEE Journal on Selected Areas in Communications*, 14, 4(May), 711 - 717.

$E(m)=np_s+(r+l)p_m/2$，方差为 $\sigma^2=np_s(1-p_s)+(r+l)p_m/2(1-p_m/2)$ 的高斯分布。过载概率就可以近似成这个高斯分布超过容量 N 的概率。如果令过载概率为最大值 q，就可以得到相应的小区中可以支持的呼叫数量 n 的最大值。这提供了上面列出的3个准入条件的第1个。具体来说，采取高斯近似，有

$$
\begin{aligned}
q &= \int_N^{\infty} \frac{e^{-[x-E(m)]^2/2\sigma^2}}{\sqrt{2\pi\sigma^2}}dx \\
&= \frac{1}{2}\mathrm{erfc}(a)
\end{aligned}
\tag{9.41}
$$

这里 erfc(a)为互补误差函数，定义为

$$
\mathrm{erfc}(a) \equiv \frac{2}{\sqrt{\pi}}\int_a^{\infty} e^{-x^2}dx \tag{9.42}
$$

参数 a 由下面给出

$$
\begin{aligned}
a &= \frac{N-E(m)}{\sqrt{2\sigma^2}} \\
&= \frac{N-(np_s+(l+r)p_m/2)}{\sqrt{2\left[np_s(1-p_s)+(l+r)\frac{p_m}{2}(1-p_m/2)\right]}}
\end{aligned}
\tag{9.43}
$$

给定最大的过载概率 q，利用式(9.41)和式(9.42)，可以找到相应的 a 的值。利用式(9.43)可以得到比较合适的 n 的值。把这个值称为 n_1，因为它代表第1个准入条件。对于条件2和条件3，重复一样的过程，得到 n_2 和 n_3。最终的准入判决就是选择小区 C_n 所支持的最大呼叫数量的3个值当中的最小值。

$$
n=\min(n_1,n_2,n_3) \tag{9.44}
$$

图9.7和图9.8展示了仿真结果，比较了这种分布式状态估计准入控制方案和保护信道的方案(Naghshineh 和 Schwartz, 1996)。将10个小区排成一个圆，构成一个一维系统，来进行这种仿真。选择保护信道的保留参数 g 的值，当使每个小区名义上的厄兰负载是 $0.85N$ 的时候，两种方案性能大致相当。平均呼叫持续时间是500 s。平均切换时间是 $1/\eta=100$ s，估计时间 T 选为20 s。在 T s中的最终切换概率是 $p_m=1-e^{-\eta T}=0.18$。随着厄兰负载的变化，图9.7比较了两种准入控制方案，每个小区有 $N=20$ 个信道。图9.8在 $N=50$ 的时候提供了相同的比较。在较高的负载时，分布式的方案提供了较低的掉话概率。把相邻小区的状态也考虑进去，在较高负载的时候，相对于使用静态的保护信道方案来说，系统资源可以得到更有效的利用。通过修改分布式的准入控制方案，甚至可以得到更好的性能(Epstein 和 Schwartz, 1998)。当一个新业务到达的时候，每个小区计算预测的掉话概率，如果3个概率中的最大值小于或者等于参数 q 的值，那么新的呼叫就会被接入。然后，我们发现切换掉话概率相对于负载几乎保持不变。这种算法被扩展到多种等级流量的情况，由于不同的等级需要不同的信

道带宽，因此具有不同的保持时间特性(Epstein 和 Schwartz，1998)[①]。还有两种分布式的多等级准入控制算法被提出来——用预测来得到一个独立于负载的切换掉话概率，同时提供一个指定的阻塞概率轮廓(Epstein 和 Schwartz，2000)[②]。还提出了其他的一些准入控制策略，可参阅 Yu 和 Leung，1997[③]；Choi 和 Shin，1998[④]；Chao 和 Chen，1997[⑤]；Levine 等，1997[⑥]的著作。Yu 和 Leung，(1997)[⑦]描述了保护信道技术的一种动态版本。Sutivong 和 Peha(1997)[⑧]提供了很多准入控制策略的比较。

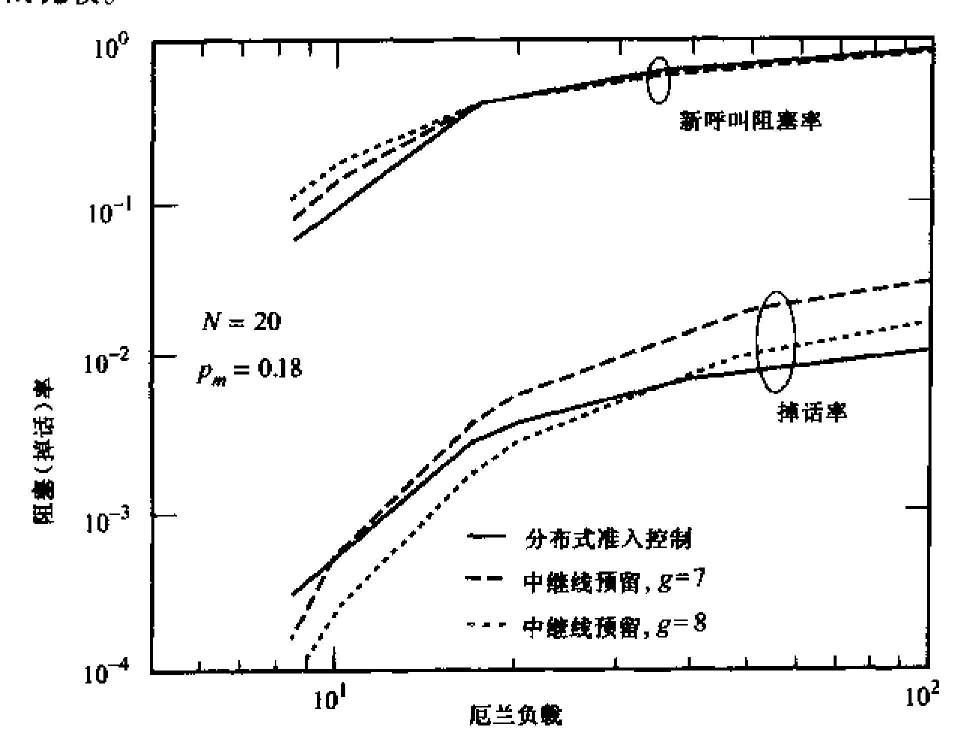


图 9.7　$N = 20$ 时采用中继线预留的方法进行状态估计的比较(引自 Naghshineh 和 Schwartz，1996，Fig.6)

下一部分我们将把工作扩展到二维蜂窝网络模型上。分析方法是类似的，但是由于几何形状的关系而更加复杂。这里得到的一些概率分布的方程是独立于特定的几何形状的，所以可以直接应用。

---

① Epstein, B. and M. Schwartz. 1998. QoS-Based Predictive Admission Control for Multi-Media Traffic, in *Broadband Wireless Communications*, ed. M. Luise and S. Pupolin, Berlin, Springer-Verlag, pp. 213 - 224.

② Epstein, B. and M. Schwartz. 2000. "Predictive QoS-based admission control for multiclass traffic in cellular wireless networks," *IEEE Journal on Selected Areas in Communications*, 18, 3 (March), 523 - 534.

③ Yu, O. and V. Leung. 1997. "Adaptive resource allocation for prioritized call admission in ATM-based wireless PCN," *IEEE Journal on Selected Areas in Communications*, 15, 9 (September), 1208 - 1225.

④ Choi, S. and K. G. Shin. 1998. "Predictive and adaptive bandwidth reservation for handoffs in QoS-sensitive cellular networks," Proc. SIGCOMM98, 155 - 166.

⑤ Chao, C. and W. Chen. 1997. "Connection admission control for mobile multiple-class personal communication networks," *IEEE Journal on Selected Areas in Comunications*, 15, 10 (October), 1618 - 1626.

⑥ Levine, D. *et al*. 1997. "A resource estimation and admission control algorithm for wireless multimeida networks using the shadow cluster concept," *IEEE/ACM Transactions on Networking*, 5, 1 (February), 1 - 12.

⑦ Sutivong, A. and J. Peha. 1997. "Call admission control algorithms: proposal and comparison," Proc. IEEE Globecom.

⑧ Guerin, R. A. 1987. "Channel occupancy time distribution in cellular radio systems," *IEEE Transactions on Vehicular Technology*, VT-35, 3 (August), 89 - 99.

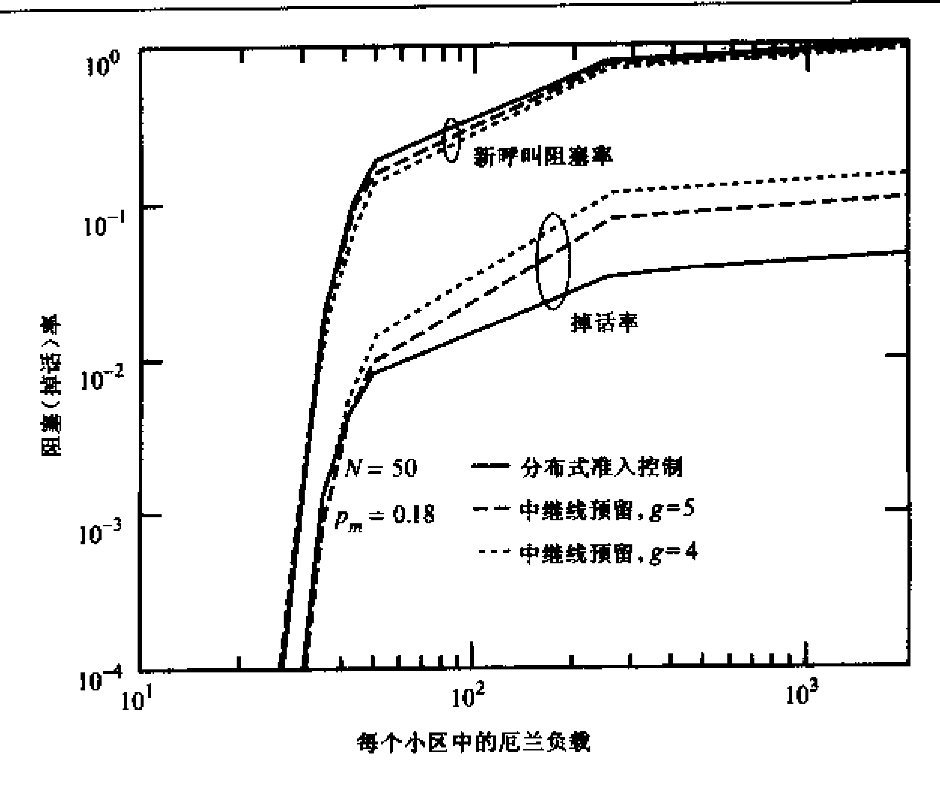


图 9.8 $N=50$ 时采用中继线预留的方法来进行状态估计比较(引自 Naghshineh 和 Schwartz，1996，Fig.6)

9.3 二维小区

现在把先前的一维分析拓展到二维小区的情况。在这样做的时候，我们会发现先前的一些分析和基本方程完全适用，与维数独立。在这里我们遵照 Hong 和 Rappaport(1999)的工作，只有对固定速率的情况的分析不太一样，这的确简化了分析。Hong 和 Rappaport 提供了对于更一般的情况的解决方案，移动速率是一个均匀分布的随机变量。然而，从先前的分析可以看出，固定速率的情况和速率服从均匀分布时得到的结果完全一致。这还将涉及 Guerin(1987)提出的工作，他在研究二维蜂窝系统的性能时使用了一个与 Hong 和 Rappaport 有点不同的模型。

这里的目标是确定信道保持时间的分布 $F_c(t)$，就像式(9.37a)中的一样。把它近似成一个等价的指数分布，平均信道保持时间是 $1/\mu_c$，就像前面式(9.38)和式(9.39)中做的那样，可以估计保护信道准入控制策略的阻塞概率和切换掉话概率，还有文献中提出的其他切换优先级准入控制策略。在前面章节里描述了一种这样的预料性过程。文献中出现的其他的预测，在前面也有所涉及。对于二维的情况，Hong 和 Rappaport 也估计了一种方案的性能，在那种方案里，如果切换到一个新的小区时信道不能立即可用，一个切换业务会被缓存起来。这个过程显然把切线掉话率折中为获取信道上的一个延时。如果延时在几分之秒量级，移动用户是不会明显地感觉到有影响的。

由式(9.37a)可知，对于新业务和切换业务而言，信道保持时间的分布 $F_c(t)$分别依赖于小区的驻留时间分布 $F_N(t)$和 $F_H(t)$。对于一维几何小区，在固定速率和均匀分布的时候，前而的章节已经计算了这两种分布。现在根据 Hong 和 Rappaport 的方法来计算二维几何，但只计算固定速率的情况。一旦确定了驻留时间分布，信道保持时间分布直接由式(9.37a)得到。前面提到，Hong 和 Rappaport 计算过随机速率的情况。为了简化计算，假定小区是圆形，如

图 9.9 所示。Guerin(1987)使用了六边形的小区，结果计算有些复杂。为了把这里得到的结果和六边形的情况联系起来，让两种情况下小区的面积相等。正如在第 6 章中使用两种小区的几何形状来比较 CSMA 的冲突计算(如图 6.6 和图 6.7 所示)。R 是六边形小区的半径，R_S 是等价的圆形小区的半径，有 $\pi R_S^2 = 3\sqrt{3}R^2/2$，或者 $R_S = 0.91R$。

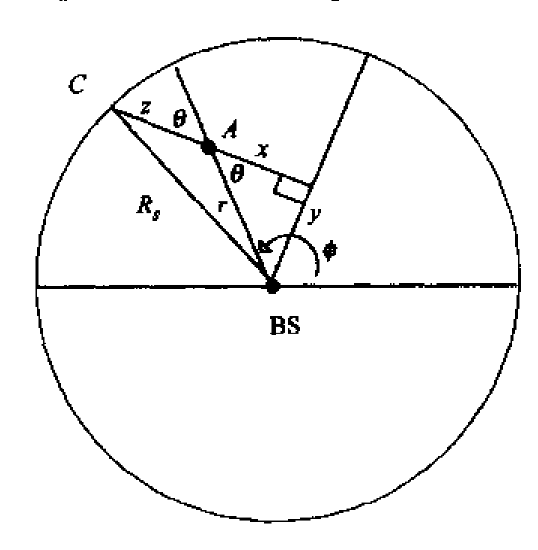


图 9.9　新的小区驻留时间的计算，圆形小区模型

先考虑新业务到达的情况。假定移动台均匀分布于小区的任何地方。假设在点 A 开始发起一个新的呼叫，如图 9.9 所示。这一点与水平轴成ϕ角，距离小区中心的基站为 r。假定移动台在点 A 以一个恒定的速率 V_c 移动。由于假定移动台均匀分布在小区中，所以角度ϕ是一个随机值，均匀分布于 0 到 2π 之间。移动台以相同的方向连续运动，到达了圆形小区边界上的点 C，与 A 的距离是 z。小区的驻留时间就是移动台走过距离 z 的时间。这个时间，$T_N = z/V_c$，是一个随机变量，要计算它的分布 $F_N(t)$。当 V_c 不变时，这个分布显然和 z 的分布成正比。这个分布 $F_z(z)$必须先计算。

按照 Hong 和 Rappaport 的方法，用图 9.9 中画出的随机的角 θ 代替ϕ，确定 θ 和 r 的联合分布，然后替换成等价的 x 和 y 的联合分布，x 和 y 两个随机变量也是如图 9.9 中定义的，然后把 z 和这些量联系起来。这里 $x = r\cos\theta$，$y = r\sin\theta$。具体说来，由于假定了移动台在整个圆形小区内呈均匀分布，θ 一定均匀分布于 0 到 π 上，而 r 的密度函数$f_r(r) = 2r/R_s^2$(显然这两个随机变量是相互独立的)。因为一个微分域事件的初等概率和相关系数的选择无关，所以不管是 x,y 还是 r,θ，一定有

$$f_{x,y}(x,y)\mathrm{d}x\mathrm{d}y = f_r(r)f_\theta(\theta)\mathrm{d}r\mathrm{d}\theta = \frac{2r}{\pi R_s^2}\mathrm{d}r\mathrm{d}\theta \tag{9.45}$$

但是 $\mathrm{d}x\mathrm{d}y = r\mathrm{d}r\mathrm{d}\theta$，因此有

$$f_{x,y}(x,y) = 2/\pi R_s^2 \tag{9.46}$$

由图 9.9 看出，$(z+x)^2 + y^2 = R_s^2$。利用这个方程，可以把$f_{x,y}(x,y)$转换成一个等价的联合密度函数 $f_{z,w}(z,w)$，w 是一个虚拟变量，从中可以很容易地得到想要的 z 的密度函数 $f_z(z)$。具体说来，令 $w = x$，那么就有 $y^2 = R_s^2 - (z+w)^2$。用两个联合分布的随机变量向另两

个随机变量进行标准的变换可得到 $f_{z,w}(z,w)$

$$f_{z,w}(z,w)=\frac{f_{x,y}(x,y)}{|J(x,y)|} \tag{9.47}$$

$J(x,y)$是变换的雅克比系数，定义为

$$J(x,y)=\begin{vmatrix}\frac{\partial z}{\partial x} & \frac{\partial z}{\partial y}\\ \frac{\partial w}{\partial x} & \frac{\partial w}{\partial y}\end{vmatrix}=\begin{vmatrix}\frac{\partial x}{\partial z} & \frac{\partial x}{\partial w}\\ \frac{\partial y}{\partial z} & \frac{\partial y}{\partial w}\end{vmatrix}^{-1} \tag{9.48}$$

这个变换就是将式(9.31)用到的一个变量的变换扩展到两个变量[上面当从 $f_r(r)f_\theta(\theta)$得到 $f_{x,y}(x,y)$的时候就已经用到了这个变换，但是过程要直接和简单得多]。

通过计算式(9.48)中的偏导，然后计算最终的行列式，读者可以证明，要求的概率密度函数 $f_{z,w}(z,w)$为

$$f_{z,w}(z,w)=\frac{|z+w|}{\sqrt{R_s^2-(z+w)^2}}\cdot\frac{2}{\pi R_s^2} \tag{9.49}$$

为了得到 $f_z(z)$，对变量 w 进行积分

$$f_z(z)=\int_w f_{z,w}(z,w)\mathrm{d}w \tag{9.50}$$

但是$(w+z)=\sqrt{(R_S^2-y^2)}$，很明显，必须有 $0\leqslant y\leqslant R_S$(见图 9.9，$y=r\sin\theta$，$r\leqslant R_S$)，因此 $0\leqslant w+z\leqslant R_S$。然而，$0\leqslant z\leqslant 2R_S$，$w=x$，$-R_S\leqslant w\leqslant R_S$。结合这些不等式，有 $-z/2\leqslant w\leqslant R_S-z$。由式(9.49)和式(9.50)，有

$$\begin{aligned}f_z(z)&=\int_{-z/2}^{R_s-z}\frac{2}{\pi R_s^2}\frac{(z+w)}{\sqrt{R_s^2-(z+w)^2}}\mathrm{d}w\\&=\frac{2}{\pi R_s^2}\sqrt{R_s^2-(z/2)^2},\quad 0\leqslant z\leqslant 2R_s\end{aligned} \tag{9.51}$$

要求的新呼叫的小区驻留时间 T_N 的分布 $T_N=z/V_c$，V_c 是速率，一个给定的常数值。一般情况下，z 和 V 都是随机变量，前面的部分已经完成了一种情况，可以再次用式(9.26)，根据式(9.51)和假定的速率分布来确定 $f_N(t)$。这些就是 Hong 和 Rappaport(1999)对正在考虑的二维蜂窝小区所做的工作。计算虽然相当直观，但是很乏味，因此选择了固定速率的情况，这样可以简化计算。基于前面对一维情况的讨论，不希望结果与固定速率的时候差别太大。虽然移动速率恒定，但是由于方向是随机的，所以在运动时还是有些不确定性。感兴趣的读者可以翻阅 Hong 和 Rappaport 的著作，里面有针对更一般的情况——移动速率均匀分布的时候的计算。

所有移动台的运动速率都是一个恒定的值 V_c，这时从式(9.51)可以得到一个新呼叫的小区驻留时间的概率密度函数

$$\begin{aligned}f_N(t)=V_cf_z(V_ct)=\frac{2V_c}{\pi R_s^2}\sqrt{R_s^2-\left(\frac{V_ct}{2}\right)^2},\\0\leqslant t\leqslant 2R_s/V_c\end{aligned} \tag{9.52}$$

定义 $K \equiv V_c/R_S$,这个表达式可以写成简单一点的归一化形式,有

$$f_N(t) = \frac{2K}{\pi}\sqrt{1-\left(\frac{Kt}{2}\right)^2}, \quad 0 \leqslant \frac{Kt}{2} \leqslant 1 \tag{9.52a}$$

积分以后,相应的概率分布 $F_N(t) = \text{Prob}(T_N \leqslant t)$为

$$F_N(t) = \int_0^t f_N(t)\mathrm{d}t = \frac{2}{\pi}\left[\frac{Kt}{2}\sqrt{1-\left(\frac{Kt}{2}\right)^2} + \sin^{-1}\left(\frac{Kt}{2}\right)\right] \tag{9.53}$$

简单一想就可以知道,$F_N(t)$从 0 到 1 上是单调变化的。特别地,对于 $Kt/2 = 1$ 即 $t = 2/K = 2R_S/V_c$,有 $F_N(t) = 1$。对于 $Kt/2 = 1/2$ 即 $t = R_S/V_c$, $F_N(t) = 0.609$。因此,一个携有新业务请求的移动台有 61%的概率会在 R_S/V_c 时间内离开小区。对于 $Kt/2 = 0.707$ 即 $t = 1.414R_S/V_c$, $F_N(t) = 0.818$。

现在考虑确定带有切换业务的移动台的驻留时间分布为 $F_H(t)$。图 9.10 画出了相关的几何图。一个移动台到达一个小区的边界点 A,以速率 V_c 和任意角度 θ 在小区中移动了距离 z,如图所示,到达了边界上的 C 点。速率仍旧是个常数值 V_c,那么驻留时间就是 z/V_c。从图 9.10 的几何关系看,从进入本小区时的第 1 个切换点 A 到进入相邻小区的第 2 个切换点 C,移动的距离 $z = 2R_S\cos\theta$。角度 θ 均匀分布于 $-\pi/2$ 到 $\pi/2$,这样,移动台就会均匀分布于整个小区。先求 $F_z(z)$,然后求 $F_H(t)$。$F_z(z)$可以直接由均匀分布的 θ 确定。

$$\begin{aligned} F_z(z) &= \text{Prob}(Z \leqslant z) = \text{Prob}\left(|\theta| \geqslant \cos^{-1}\frac{z}{2R_s}\right) \\ &= 1 - \frac{2}{\pi}\cos^{-1}(z/2R_s) \quad 0 \leqslant z \leqslant R_s \end{aligned} \tag{9.54}$$

这个分布也可以写成下面的等价形式:

$$F_z(z) = \frac{2}{\pi}\sin^{-1}(z/2R_s) \tag{9.54a}$$

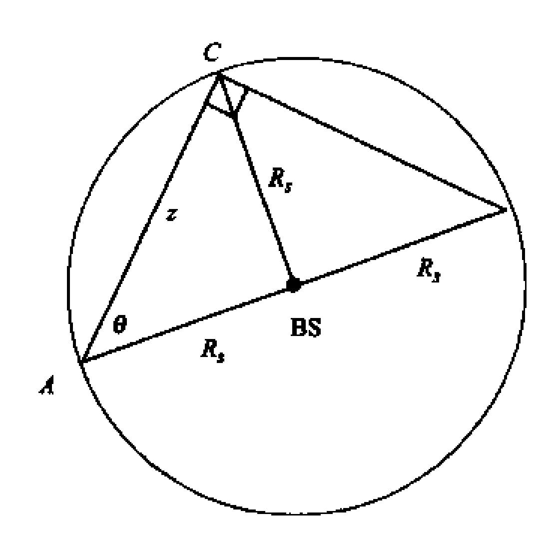


图 9.10 切换呼叫驻留时间的计算

概率密度函数 $f_z(z)$就是 $F_z(z)$的导数

$$f_z(z)=\frac{1}{\pi R_s}\frac{1}{\sqrt{1-(z/2R_s)^2}} \tag{9.55}$$

由于假定速率 V 是一个恒定的值 V_c, $t=z/V_c$,立即有 $f_H(t)=V_cf_z(V_ct)$,即

$$f_H(t)=\frac{V_c}{\pi R_s}\frac{1}{\sqrt{1-(V_ct/2R_s)^2}} \tag{9.56}$$

定义 K 为 V_C/R_S,这个切换驻留时间的密度函数 $f_H(t)$也可以写成归一化的形式

$$f_H(t)=\frac{K}{\pi}\frac{1}{\sqrt{1-(Kt/2)^2}} \qquad K\equiv V_c/R_s \tag{9.56a}$$

分布函数 $F_H(t)$为

$$F_H(t)=\int_0^t f_H(t)\mathrm{d}t=\frac{2}{\pi}\sin^{-1}\left(\frac{Kt}{2}\right) \tag{9.57}$$

[将结果和式(9.54a)中的 $F_z(z)$进行比较,实际上,由于 $t=z/V_c$,可根据式(9.54a)在固定速率的情况下直接写出来]。注意,比较式(9.57)和式(9.53),也可以写出

$$F_N(t)=F_H(t)+\frac{2}{\pi}\frac{Kt}{2}\sqrt{1-\left(\frac{Kt}{2}\right)^2} \tag{9.53a}$$

知道了两个小区驻留时间的分布 $F_N(t)$和 $F_H(t)$,利用式(9.37a),可以计算信道保持时间的分布 $F_c(t)$,正如前面计算一维的情况一样。把这些结果同等价的指数分布做比较,看看对于信道保持时间,指数近似的效果如何。这是在分析保护信道准入控制方案时采取的假设。但是首先计算两种切换概率 P_N 和 P_H,把二维的结果同一维的结果进行比较。为了这个目的,再次使用式(9.19)和式(9.20)。先考虑新业务的切换概率 P_N。根据式(9.19)和式(9.52),有

$$P_N=\int_0^\infty \mathrm{e}^{-\mu t}f_N(t)\mathrm{d}t=\frac{4}{\pi}\int_0^1\sqrt{1-x^2}\mathrm{e}^{-bx}\mathrm{d}x \qquad b\equiv 2\mu R_s/V_c \tag{9.58}$$

标量参数 b 和前面一维小区模型中出现的标量参数 $a\equiv\mu L/V_c$ 是可比的。实际上,令二维的直径 $2R_S$ 和一维的小区长度 L 相等,然后比较一维和二维的例子。

一个移动台携带着正在进行的通话到达一个小区,可由式(9.20)和式(9.56)得到切换的概率 P_H

$$P_H=\int_0^\infty \mathrm{e}^{-\mu t}f_H(t)\mathrm{d}t=\frac{2}{\pi}\int_0^1\frac{\mathrm{e}^{-bx}}{\sqrt{1-x^2}}\mathrm{d}x \tag{9.59}$$

标量参数 $b\equiv 2\mu R_S/V_c$,正如上面写 P_N 的表达式时定义的那样。表 9.3 中计算切换概率的例子和前面表 9.2 中的 3 个例子是一样的。我们提供的还包括表 9.2 中一维情况下的切换概率,并对一维和二维的情况做了一个比较。列出来做比较的还有 γ_c 的值,它定义为当阻塞概率和切换掉话率都很小时,切换业务 λ_h 和新到业务 λ_n 的比值。由式(9.22)可知,该比值正是

$P_N/(1-P_H)$,它是用来比较结果的惟一参数。γ_c 的这些值在计算信道保持时间的分布 $F_c(t)$ 的时候也要用到,这从式(9.37a)中很容易看出。正如上面提到的那样,在比较的时候令一维的小区长度 L 和蜂窝小区的直径 $2R_S$ 相等。因此,这里引入的参数 b 和一维情况下引入的参数 $a\equiv\mu L/V_c$ 相等(在做这些比较的时候只选择固定速率模型)。在所有的3个例子中,平均呼叫保持时间 $1/\mu$ 还是取 200 s。

表 9.3　切换概率

一维情况			二维情况		
P_N	P_H	γ_c	P_N	P_H	γ_c
1 宏小区, $2R_S=L=10$ km, $V_c=60$ km/h, $b=3$					
0.32	0.05	0.34	0.37	0.23	0.48
2 宏小区, $2R_S=L=1$ km, $V_c=60$ km/h, $b=0.3$					
0.86	0.74	3.31	0.88	0.83	5.18
3 微小区, $2R_S=L=100$ m, $V_c=5$ km/h, $b=0.36$					
0.84	0.70	2.8	0.86	0.80	4.3

尽管在所有的3个例子中 $\gamma_c|_{一维}<\gamma_c|_{二维}$,一维和二维的例子的结果还是相当的。在相同的移动速率的情况下,较小的宏小区的切换概率比大一些的宏小区的切换概率高得多。这些结果在预料之中,评论和前面的一样。很容易证明,如果选二维圆形小区的半径 R_S 等于可比的一维小区的长度 L,而不是 L 的 1/2,则得到的 γ_c 一律比二维情况下的 γ_c 小。因此,一维情况下的结果被限定在二维的结果中。稍后,在比较二维蜂窝小区模型的结果和 Guerin(1987)的著作中出现的结果时,还要回到这些 γ_c 的计算上来,后者使用了六边形的小区结构,建模方法也有些不同。

现在计算信道保持时间的分布 $F_c(t)$ 和它的具有相同均值的等价指数分布的近似值。上面曾经反复说过,指数近似对于计算保护信道准入控制方案的性能是必需的。如果被证明是足够精确的话,那么使用它也可以简化其他的性能计算。假定速率固定,由式(9.37a)、式(9.53)和式(9.57)直接可写出二维圆形小区模型的 $F_c(t)$ 表达式

$$F_c(t)=1-e^{-\mu t}\left[1-\frac{2}{\pi}\sin^{-1}(Kt/2)-\frac{Kt}{\pi(1+\gamma_c)}\sqrt{1-\left(\frac{Kt}{2}\right)^2}\right] \tag{9.60}$$

回忆一下,$K\equiv V_c/R_S$,V_c 是所有终端的速率,假设是恒定的,R_S 是小区的半径。正如上面提到的,参数 γ_c 是切换业务与新到业务的比值,对于阻塞概率和切换掉话概率很小的情况,它就是 $P_N/(1-P_H)$。注意,按照要求,式(9.60)有时间 t 上的限制,$0\leqslant t\leqslant 2/K=2R_S/V_c$。定义一个变量 $x\equiv Kt/2=V_ct/2R_S$,式(9.60)可以写成归一化的形式,更适合画图。然后,我们得到 $F_c(x)$

$$F_c(x)=1-e^{-bx}\left[1-\frac{2}{\pi}\left(\sin^{-1}x+\frac{x\sqrt{1-x^2}}{1+\gamma_c}\right)\right]\qquad 0\leqslant x\leqslant 1 \tag{9.60a}$$

参数 b 就是早些时候定义的量 $2\mu R_S/V_c$。

现在用式(9.39)计算平均信道保持时间 $1/\mu_c$，我们在近似指数信道保持时间分布时会用到它。利用式(9.60a)给出的 $F_c(t)$归一化形式，有

$$\frac{1}{\mu_c}=\frac{1}{\mu}\left[1-\mathrm{e}^{-b}-\frac{2b}{\pi}\int_0^1 \mathrm{e}^{-bx}\left(\sin^{-1}x+\frac{x\sqrt{1-x^2}}{1+\gamma_0}\right)\mathrm{d}x\right] \tag{9.61}$$

作为计算 $1/\mu_c$ 的例子，用式(9.61)考虑表 9.3 中 3 个二维的情况。对于第 1 种情况，宏小区半径 5 km，移动速率 $V_c=60$ km/h，$b=3$，$1/\mu_c=0.675/\mu=135$ s。对于宏小区半径为 0.5 km 的第 2 种情况，具有相同的移动速率 60 km/h，$b=0.3$，$1/\mu_c=0.16/\mu=32$ s。这些结果和在前面的那些一维的例子中得到的结果一致。随着小区半径的减小，移动速率保持不变，切换的概率显著增加(见表 9.3)，平均信道连接时间迅速减少。

对于二维的圆形小区模型，在平均连接时间 $1/\mu_c$[由式(9.61)给出]相同的时候，指数连接时间的近似与实际分布概率 $F_c(t)$[由式(9.60)或式(9.60a)给出]有多接近？分别比较表 9.3 中例 1 和例 2 在图 9.11 和图 9.12 中的两个分布。在比较 $F_c(x)$和 $1-\exp(bx\mu_c/\mu)$时，使用分布的归一化版本。注意，对于 $b>1$，也就是第 1 个例子的情况，指数近似是非常准确的。第 2 个例子的结果，也就是较小的小区，$b=0.3$，是非常差的。通过比较方程，原因很明显。速率固定的模型导致信道保持时间最大是 $2R_S/V_c$，或者是归一化的形式 $x=1$。指数近似就没有这样的界限。对于相对大的小区，速率不是太高，指数近似和用模型得到的分布都以足够快的速率趋向于极限概率 1，因此分布在时间上受限而指数近似不受限的事实不会对比较影响得太厉害。当 $b=0.36$ 时，例 3 的结果和第 2 个例子非常像，比较非常差。

现在考虑其他两个参数 $b>1$ 的例子。这些是 Guerin(1987)的著作中考虑六边形小区的例子，我们将比较它们的结果和这里得到的分析结果。对于第 1 个例子，把圆形小区近似成一个半径为 $R=1$ mile的六边形小区；移动速率取每小时 35 mile，而平均呼叫保持时间取 2 min。等价的圆形小区半径 R_S 为 0.91 mile，参数 $b=1.56$。图 9.13 比较了两个信道的保持时间分布，一个用式(9.60a)表示，另一个是带有相同平均值 $1/\mu_c$ 的指数近似。指数近似是非常好的。在图 9.14 的情况下，小区半径大了 10 倍，速率和呼叫保持时间保持不变，这种近似甚至更好。因此，对于 $b>1$ 的情况，指数近似非常好。这与 Guerin(1987)直接用六边形模型得到的关于信道保持时间分布的结果是相吻合的，而不是迄今讨论的圆形小区模型。

Guerin 采用的建模方法可以总结如下：假定移动台都具有相同的固定速率 V_c，但是在呼叫初始化的时候朝 4 个直角方向运动的概率是一样的。一旦在通话开始时选定了方向，则在通话过程中保持方向不变。然后，通过直接的计算详细分析了一个二维的六边形系统的信道保持时间分布。分析和仿真证实了指数分布是信道保持时间分布的一个极好的近似，尤其是在小区尺寸较大的情况下，正是上面用圆形小区模型得到的结果。Guerin 用到的指数近似的每个呼叫切换的平均数量 $E(h)$是用六边形模型计算出来的。有趣的是，在两个分布中，使每个呼叫的平均切换数相等，与使平均连接保持时间 $1/\mu_c$ 相等，被证明是等价的，它们分别是本节和前面采用的分析方法。用两种方法可证明这种等价性。

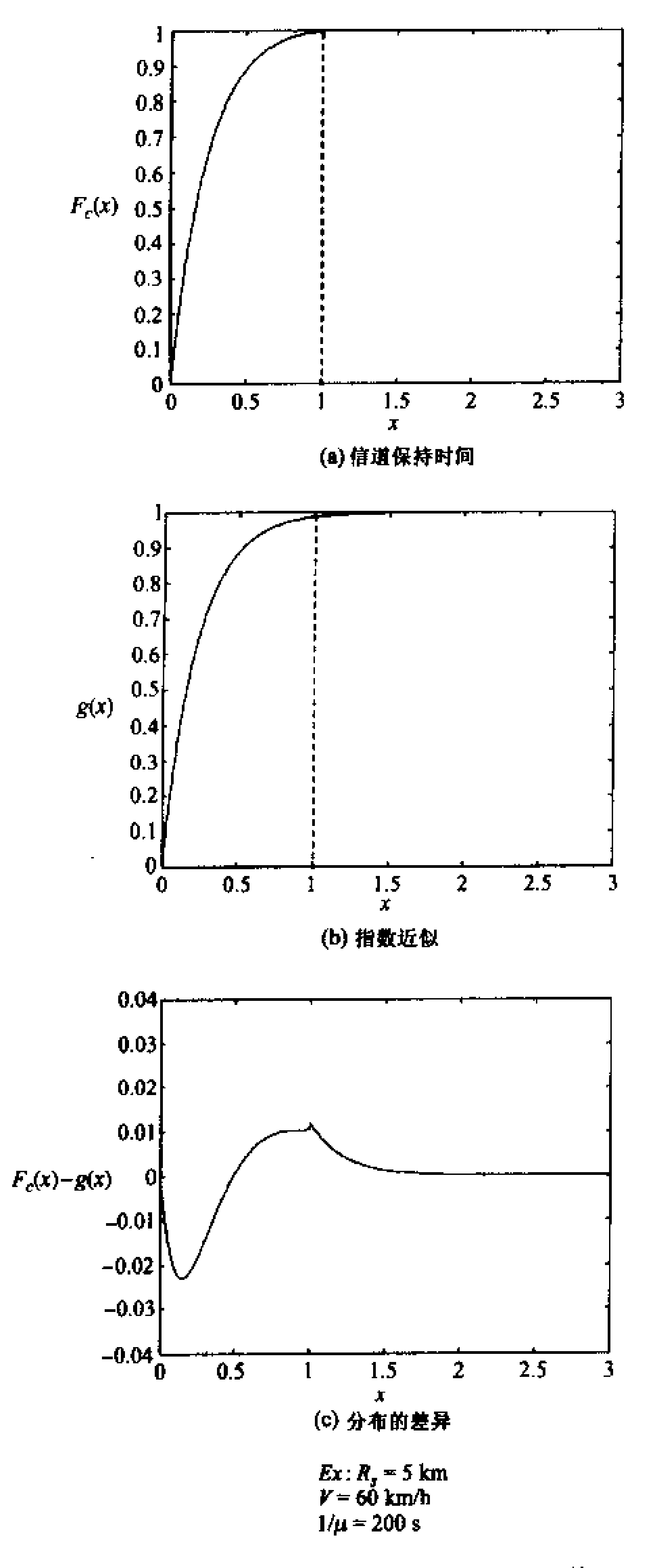


图 9.11　指数近似的信道保持时间分布的比较，表 9.3 中的情况 1，$b = 3$

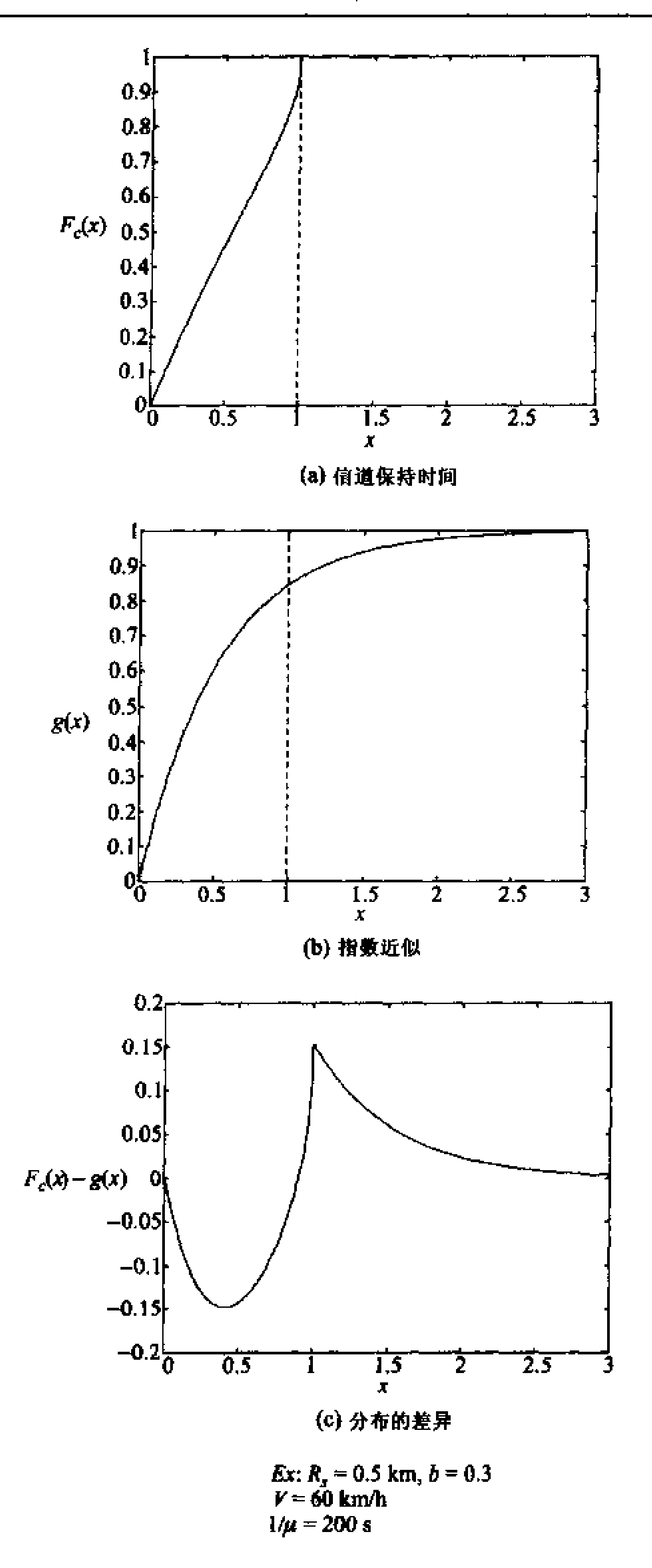


图 9.12 指数近似的信道保持时间分布的比较，表 9.3 中的情况 2，$b=0.3$

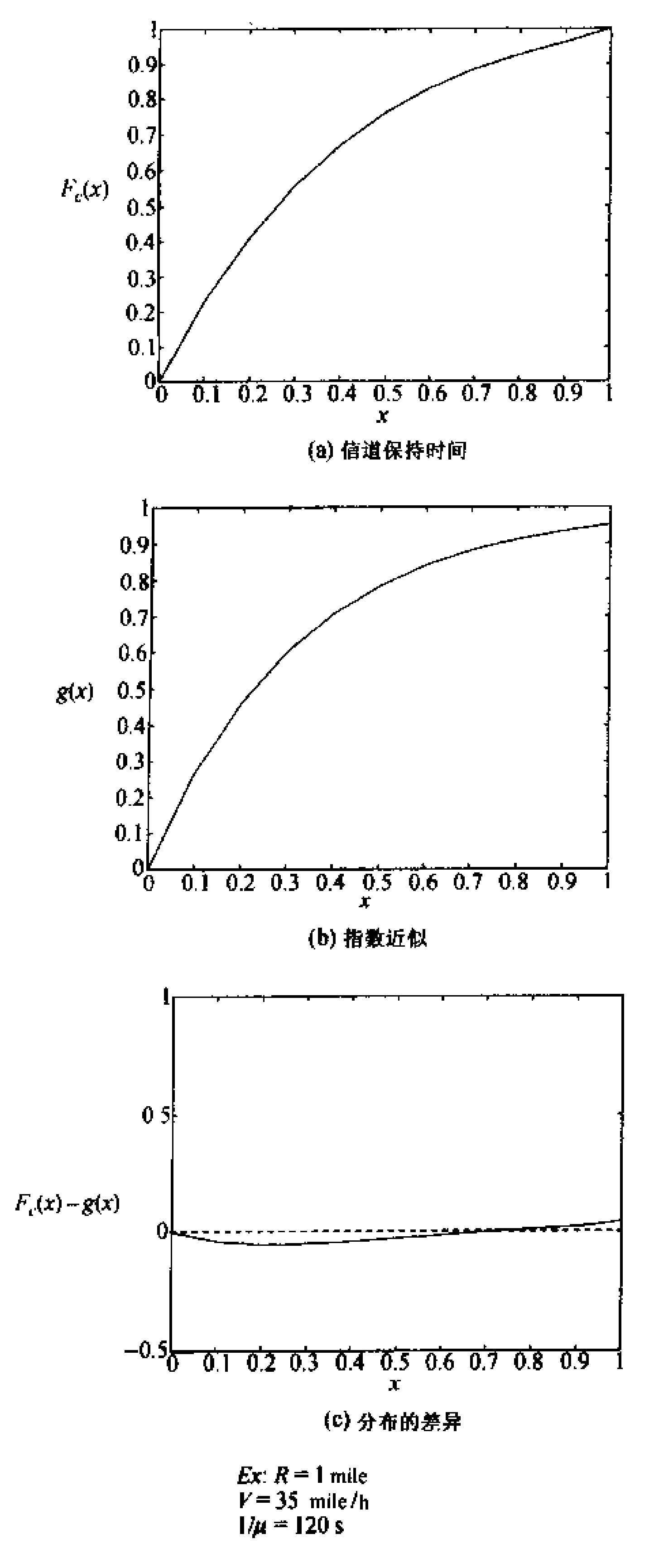


图 9.13　指数近似的信道保持时间分布的比较，$b = 1.56$

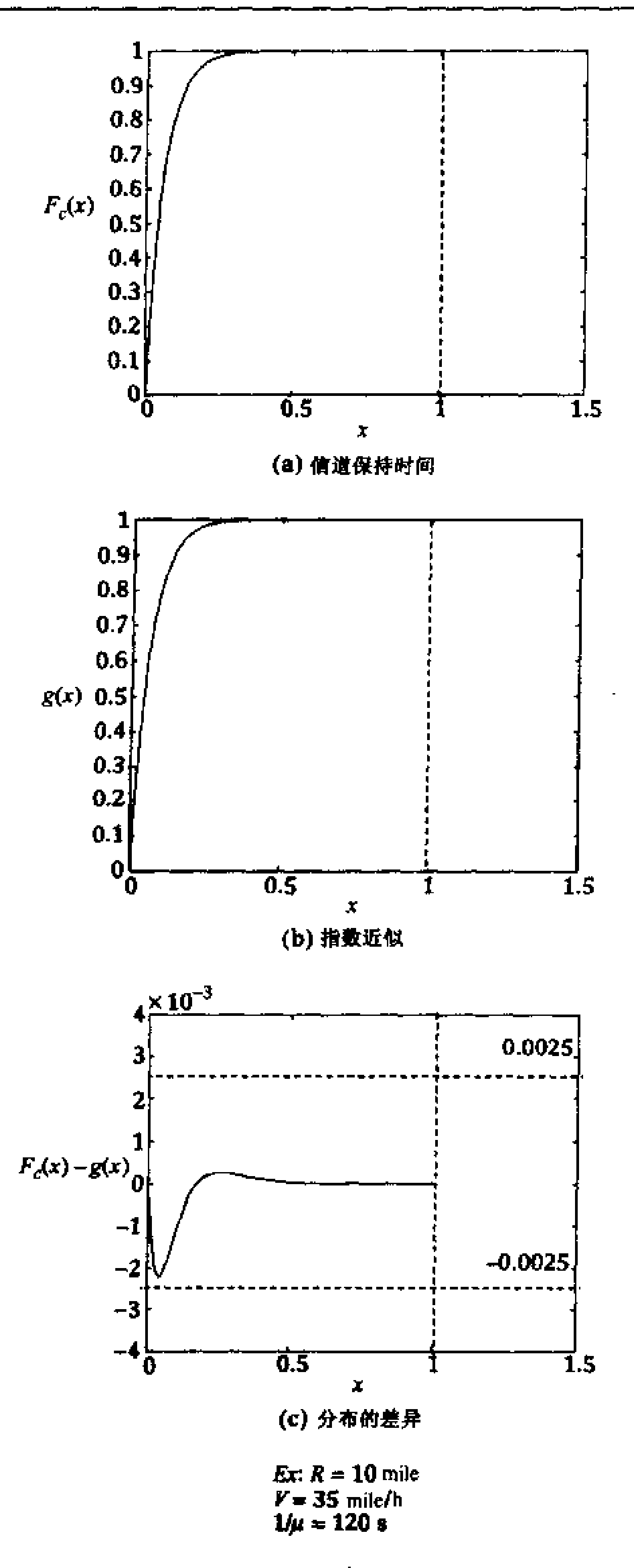


图 9.14 分布的比较，$b = 15.6$

对于六边形蜂窝移动系统的详细分析，证明了每个呼叫切换的平均数量是下面简单的表达式：

$$E(h)=(3+2\sqrt{3})/9\alpha=0.72/\alpha,\ \alpha\equiv R\mu/V_c,\ R \text{ 是六边形的半径}$$

但是，注意这个贯穿于 Guerin 的分析的归一化参数 α 和在圆形小区模型中使用的参数 $b\equiv 2R_S\mu/V_c$ 是直接成正比的。实际上，由于等价于六边形的圆形小区的 $R_S=0.91R$，所以有 $b=1.82\alpha$，或者使用符号 $E(h)=1.3/b$。但我们用例子证明了对于圆形小区模型，如果 $b>1$，指数近似是非常好的。这表明每个呼叫切换的平均数量的阶数是 1 或者更小。Guerin 也得到了同样的结论。这样的一个系统需要一个足够大的小区，不太高的速率，或者一个充分小的呼叫保持时间，这样就不会发生太多的切换。所有的这些都被一个重要的参数 $E(h)$ 联系在一起。

通过简单地计算，Guerin 发现信道保持时间取决于简单参数 α，那么对于圆形小区，一个类似的结果是不是就不成立了呢？答案证明是肯定的。考虑在圆形小区中每个呼叫的平均切换数量 $E(h)$ 的计算。无论什么几何形状，这个量一定是切换业务和信道业务的比值。这已被式(9.7b)所证实，其中忽略了阻塞概率和切换掉话概率。那里出现的比值 η/μ 正是每个呼叫切换的平均数量 $E(h)$。而根据式(9.22)，在同样假定阻塞概率和切换掉话率很小的情况下，$\lambda_h/\lambda_n=P_N/(1-P_H)\equiv\gamma_c$。因此 $E(h)=\gamma_c$，这个参数在比较系统性能的时候很有用。在讨论表 9.3 的 3 个例子的时候特别指出了这一点。然而，根据表 9.3，对于二维圆形小区的例子，$b\gamma_c\approx 1.5$。因此，在圆形小区的情况，$E(h)\approx 1.5/b$。六边形小区的这个值是 $1.3/b$，Guerin 进行了比较。考虑到六边形几何的 $E(h)$ 计算量是非常复杂和乏味的，而对于圆形的情况，利用封闭形式的式(9.58)和式(9.59)分别计算 P_N 和 P_M，整个过程非常简单，但 15% 的差异是非常显著的。实际上，可以很容易证明，在 $b\ll 1$ 和 $b\gg 1$ 两种极端情况下，若 $b\ll 1$，γ_c 趋近于 $\pi/2b=1.57/b$；若 $b\gg 1$，γ_c 趋近于 $4/\pi b=1.27/b$。因此，对于 γ_c 的两个计算，一个基于圆形小区模型，另一个基于六边形小区模型，当 $b>1$ 时，产生更相近的结果。表 9.4 的第 2 列和第 3 列比较了 γ_c 的值，先直接由 $P_N/(1-P_H)$ 给出，然后用 $1.5/b$ 近似，这里和表 9.3 的例子是一样的。我们可以看出近似程度是很好的。

表 9.4　指数近似和分析的比较

b	γ_c	$1.5/b$	$1/(1+\gamma_c)$	$1/(1+(1.5/b))$	$1/\mu_c/1/\mu$ [见式(9.61)]
3	0.48	0.5	0.68	0.67	0.675
0.3	5.18	5	0.16	0.17	0.16
0.36	4.3	4.17	0.19	0.19	0.188

现在转到 Guerin 采取的信道保持时间分布的指数近似上来。在那篇论文里，指数近似定义为这样一个模型，它同六边形几何和选择的移动模型计算出来的分布具有相同的平均切换数量 $E(h)$。现在证明这个方法在本质上等价于在本节和前一节采用的方法，后者要求平均信道保持时间相同。根据 9.1 节的式(9.3)，在小区驻留时间或者切换统计特性服从指数分布的情况下，平均信道保持或者占有时间 $1/\mu_c$，就是 $1/\mu_c=(1/\mu)/(1+\eta/\mu)$，其中 $1/\eta$ 是平均驻留或者切换时间，$1/\mu$ 是平均通话长度或者呼叫保持时间。上面提到过，比值 η/μ 正是每个呼叫的切换平均数量 $E(h)$，因此我们可得到 $1/\mu_c=(1/\mu)/[1+E(h)]$。Guerin 使用了 $E(h)=$

$0.72/\alpha$ 的近似,这个值是通过分析六边形小区得到的。他发现,当 $E(h)$ 的阶等于或者小于 1 的时候,这个近似和计算的信道保持时间分布拟和得很好,正如前面提到的那样。

我们还可以用对于 $1/\mu_c$ 而言简单的结果与本节的分析做比较。因为知道 $E(h)=\gamma_c$,所以我们可以写出 $1/\mu_c=(1/\mu)/(1+\gamma_c)$。这提供了表示信道保持时间的另一种方法。这个针对 $1/\mu_c$ 的表达式和用式(9.61)得到的相比会怎么样呢?结果表明,通过这个简单表达式可以很精确地近似式(9.61)。这在表 9.4 中已经表示出来了,和表 9.3 中的例子一样。除此之外,在表 9.4 中我们还列出了计算 $1/\mu_c$ 的近似值的结果,用 $1.5/b$ 代替 γ_c。将 $1/\mu_c$ 全部的 3 个值归一化为平均呼叫保持时间 $1/\mu$,是很接近的。因此,不仅指数分布对信道保持时间分布近似得很好,尤其是在 $b>1$ 的情况下——这已经被圆形小区和六边形小区模型所证实,而且平均信道保持时间 $1/\mu_c$ 用简单的表达式 $1/(1+1.5/b)$ 近似也很准确!因此,Guerin 采取的方法(把等价的指数信道保持时间分布定义成和所选模型具有相同的单个呼叫切换平均值)和本节采用的方法(让两个分布的平均信道保持时间相等)是等价的。

习题

提示:要解决下面的有些题目需要用到数学软件包。任何一款这样的软件包一般都能满足。

9.1 用自己的话解释和区别几个概念:小区驻留时间、呼叫保持时间和信道保持时间。

9.2 (a) 式(9.2)联系了几个概率分布:信道保持时间、小区驻留时间和呼叫保持时间(通话长度)。证明:对于指数分布的驻留时间和呼叫保持时间,信道保持时间也是指数分布的,平均信道保持时间和另两个时间的平均值的关系见式(9.3)。

(b) 证实在式(9.4)给出的指数分布的随机变量的情况下的切换概率 P_h。

(c) 证实和讨论式(9.5a)下面的不同小区尺寸和移动速率的 3 个例子的结果。选择其他的例子,对于每个例子,计算相同的 3 个参数——切换之间的平均时间、切换的概率和平均信道保持时间。

9.3 利用图 9.1,证明流量平衡方程[式(9.6)]是如何推导的。指数分布随机变量的假设用在推导的什么地方?[见习题 9.6(a)]。

9.4 考虑保护信道准入控制过程一段的讨论。

(a) 用图 9.3 的本地平衡的概念,证实式(9.10)和式(9.12)给出的关于状态 p_1 和 p_2 的概率的表达式。证明式(9.13)和式(9.14)给出了图 9.3 中两个区域的状态概率的一般表达式。

(b) 证明式(9.16)和式(9.17)分别给出了切换掉话率和新业务的阻塞概率。

(c) 表 9.1 给出了式(9.16)和式(9.17)迭代的结果,就像文中解释的那样,得到了在一个宏蜂窝系统和一个微蜂窝系统的例子中的切换掉话率和新业务阻塞概率。在这些例子中,一个宏小区半径是 10 km,一个微小区半径是 100 m。移动速率在宏小区中是60 km/s,微小区中是 5 km/s。每种情况下可用的信道数量是 $m=10$。对每个系统进行迭代,开始时把两个概率都设为 0,先是有 $g=1$ 个保护信道,然后把 g 设成两个。在进行的过程中用很多新业务到达速率的值,就像表 9.1 中指出的那样。证实和讨论表 9.1 中合适的项。特别地,证明保护信道的引入是怎样以阻塞概率的上升为代价来降低切换掉话概率的。

(d) 对于 9.2(c)中用到的其他例子重复(c)的计算。

9.5 (a) 证明对于阻塞概率的厄兰 B 公式由式(9.18)给出。

(b) 对于 $m=1,5,10$ 和 20 个信道,画出厄兰 B 阻塞概率和 ρ(或者 A,以厄兰为单位,正如第 3 章定义的那样)的关系。

(c) 对于计算式(9.18)的厄兰 B 阻塞概率,有个很有用的递归关系

$$\frac{1}{P_B(m)}=1+\frac{m}{\rho P_B(m-1)},\quad P_B(0)=1$$

推导这个关系式并用它计算和画出 $P_B(m)$,对于 $1\leq m\leq 20$ 和 1 厄兰 $\leq\rho\leq 20$ 厄兰。

9.6 这道习题是关于在小区驻留时间不是指数分布的情况下呼叫的切换。

(a) 为什么一般有不同的小区驻留时间,一种是对新产生的业务,另一种是对从另一个小区切换过来的业务?(提示:在小区中,哪里可以产生一个新业务?哪里产生了一个切换业务?)为什么指数分布的无记忆特性导致单一的驻留时间分布?

(b) 用图 9.1 和两种类型的切换概率的定义,新到达业务是 P_N,从另一个小区切换过来的业务是 P_H,来推导流量平衡方程[式(9.21)]。比较这个方程和假定指数分布小区驻留时间得到的那个[参见习题 9.3 和式(9.6)]。

(c) 解释一下为什么这两个切换概率由 $P_N=\text{Prob}[T_n>T_N]$ 和 $P_H=\text{Prob}[T_n>T_H]$ 给出。

(d) 从式 9.4 和上面(c)中 P_N 和 P_H 的定义开始,证明一般来说,这两个概率分别由式(9.19)和式(9.20)给出。

9.7 考虑图 9.4 中一维小区的情况,所有的终端以相同的速率 V_c 移动,朝各个方向的可能性相同。

(a) 证明新业务小区驻留时间密度函数、分布函数和切换概率分别由式(9.23)、式(9.24)和式(9.25)给出。

(b) 看原文考虑的 3 个例子:长度 $L=10$ km,移动台的移动速率为 60 km/h 的宏小区;长度 $L=1$ km,移动速率相同的宏小区;长度 $L=100$ m,移动速率是 5 km/h 的微小区。呼叫都是指数分布的,平均值是 200 s。证明新业务切换概率分别是 0.316,0.86 和 0.84(见表 9.2)。这些数字和直觉吻合吗?在速率固定的情况下新业务的切换概率取决于标量参数 $\mu L/V_c$,$1/\mu$ 是平均通话长度(呼叫保持时间),上面的数字与这个发现相一致吗?改变 3 个参数中的 1 个或多个,对其他的一些例子重新计算。

9.8 考虑长度为 L 的一维小区,如图 9.4 所示。终端的移动速率服从均匀分布,最大值是 V_m。专注于新呼叫的驻留时间和新业务的切换概率。

(a) 证明这种情况下的驻留时间密度函数,概率分布和切换概率 P_N 分别由式(9.27),式(9.28)和式(9.29)给出。

(b) 对于习题 9.7(b)中的 3 种情况,重复 P_N 的计算,在所有的情况下都令 $V_m=2V_c$。在所有的情况下,平均通话长度都是 200 s。证明得到的结果和表 9.2 相吻合:$P_N=$ 0.30,0.78 和 0.76。证明这些结果是怎样与在标量参数 $a\equiv\mu L/V_m$ 的依赖性上所期望的那些结果一致的?选择其他的例子重复计算,把得到的结果和基于这个参数的变化所期望的结果做比较。

9.9 本习题针对的是习题9.7和习题9.8中一维小区模型的例子，但是专注于驻留时间分布和从其他小区切换过来的业务的切换概率 P_H。

(a) 证明：当移动台都以相同的速率 V_c 移动时，P_H 由式(9.30)给出。

(b) 证明：如果移动台移动速率服从均匀分布，最大值是 V_m，那么驻留时间分布的密度函数就由式(9.31)给出，而切换概率由式(9.32)给出。

(c) 用习题9.7(b)中的3个例子得到每种情况的切换概率。平均通话时间是200 s。证明得到了表9.2中的结果。选择习题9.7(b)和习题9.8(b)中的其他例子，指出得到的结果在随机速率的情况下是如何验证对参数 $a \equiv \mu L/V_m$ 的依赖性的。

9.10 考虑习题9.7(b)中3个一维小区的例子。找到每种情况下，切换业务到达速率和新业务产生速率的比值。在计算的时候忽略阻塞概率和切换掉话概率。把结果与原文提供的做比较。

9.11 原文描述了当一个小区的移动驻留时间不服从指数分布的时候，确定信道保持时间概率分布的过程。这是用式(9.33)到式(9.37a)描述的。

(a) 根据文中所述推导式(9.37a)。注意，为了得到这个结果，通话时间必须假定是指数分布。

(b) 专注于文中讨论的一维小区模型。信道保持时间被近似成一个具有相同均值的指数分布。这使得式(9.39)成为逼近平均值的方程。对于习题9.7(b)中的3个例子，估计这个表达式。假定所有的移动台都是以相同的恒定速率移动的，速率在习题9.7(b)中已经指出。把得到的近似平均信道保持时间和通过假定直接得到的指数分布信道保持时间做比较。

(c) 重复习题9.4(c)中保护信道的计算，假定一维小区，使用上面(b)中得到的近似平均信道保持时间。把结果和习题9.4(c)中得到的结果做比较(参看表9.1中合适的项)。

9.12 这个习题专注于9.3节讨论的二维小区的几何形状。

(a) 考虑图9.9的圆形小区几何。写出原文列出的计算的细节，证明新业务的小区驻留时间的概率密度函数由式(9.52)和式(9.52a)给出，而分布函数由式(9.53)给出。注意，所有的移动台都是以相同的固定速率 V_c 移动的。

(b) 利用图9.10指出的几何形状，证明：根据原文的分析，切换业务驻留时间分布的概率密度函数由式(9.56a)给出，分布函数由式(9.57)给出。仍假定移动台以不变的速率移动。

(c) 对于表9.3的3种情况，计算新业务的切换概率 P_N 和切换业务再次切换的概率 P_H，将那里给出的结果进行比较。

9.13 (a) 对于表9.3中的3种二维的例子，用式(9.61)来计算近似的平均信道保持时间。

(b) 对于表9.3中的3个例子，画出信道保持时间分布式(9.60)或者式(9.60a)。如果可能的话，用(a)的结果添加近似指数分布并加以比较。对于其他的归一化参数 $b>1$ 的例子也这么做。随着 b 的增加，近似是不是像文中所说的那样有所改善呢？对于你的例子，你能解释为什么近似会有所改进吗？

第10章　2.5G/3G无线移动系统:分组数据交换

10.1　简介

本书前面几章主要介绍了用于无线电话业务(如语音呼叫、与有线网络的连接等)的第二代无线系统(2G)。本章将了解到,为了面对因特网和其他数据业务的挑战,许多人都在努力发展和寻求更先进的蜂窝网络,设计以提供更高速率的无线数据业务。无线网络的目标是传输语音、视频、图像和数据文件等多媒体业务,并且和有线网络做到无缝端对端的通信。有了这种承载数据业务的蜂窝网络,无线通信系统就由2G(第二代)发展成为2.5G或3G(第三代)。同时,更高比特率的无线局域网(WLAN)也用于提供类似的服务,并且已经在商业和学术机构开始使用。这些将在第12章讨论。新一代蜂窝网的语音通信和第二代移动通信一样,都是采用电路交换,而数据业务采用分组交换的形式,数据率高于当前正在使用的2G系统。第12章讨论的无线局域网(WLAN)专门使用分组交换。我们将接着学习第四代移动通信。第四代移动通信系统将全部采用分组交换,并且和有线网络(如因特网)的分组交换实现无缝连接。虽然语音传输会继续在3G中扮演重要角色,但是本章会重点讲述3G系统中的分组交换技术。

我们一直使用两个术语"电路交换"(circuit switching)和"分组交换"(packet switching)。电路交换正广泛应用于无线和有线的电话网络中。"电路"的意思是,当呼叫时专门有一条端对端的线路连接,该连接将保持到通话结束。在电路的有线网络部分,电路携带重复的TDM帧结构,每一帧由固定数量的时隙组成,这和GSM与IS-136比较相似。每个时隙分配给特定的电话。因此,电路路径由一系列特定的时隙组成。在CDMA系统中(例如IS-95),电路路径的一部分是移动台的无线接口部分,用于接收特定码字。而电路路径的其他部分则是有线部分,仍然使用特定的TDM时隙,这与TDMA系统、IS-136和GSM系统是一样的。

另外一种技术——分组交换应用于因特网和其他数据业务网络。把数据打成由比特组成的数据块称为包。这些包的长度可以固定也可以不固定,数据包在网络中是从源端到目的端分别传输的。从不同数据用户来的包共享一条链路。因此,在给定时间内的TDM帧并不是分配给某一固定用户。在分组交换的CDMA中,码字只是在需要传输某个包时才分配给那个包。

注意到,新的2.5G/3G蜂窝系统将使用高比特率的基于分组的数据传输。与语音通信相比,数据传输能力在2G系统中也是常用的。例如,GSM用电路交换提供数据传输的短消息业务(SMS)。这项业务在欧洲被广泛使用,但是该业务所能提供的数据比特率比较低。另外一个例子,IBM发明的CDPD是一种多层移动网络数据服务,在美国和IS-136一起使用。当一个IS-136帧中的某些时隙未被使用时,该时隙就用传输的分组包来填充。CDPD的接受能力是有限的。这项业务和短消息业务一样,也只能提供比特率传输相对较低的服务。这两项业务和当今增长迅速的因特网Web浏览相比,还不能满足宽带多媒体的需要。

多媒体因特网业务,尤其是那些和Web浏览有关的业务,包括用分组包的形式高速率传输文件、视频、图像,通常是从服务器端发送到请求的客户端。而在移动通信中,客户通常是某种类型的移动数据终端,一般是从基站发送高速率的下行数据到移动台。因特网数据传输有

两点与第二代移动通信不同:数据是以包的形式产生的,并且通常有很高的带宽或者比特率,不同的需求导致 WLAN 和 3G 各自得到发展。

3G 蜂窝系统和 WLAN 需要基于分组的更高速率的数据通信有多方面的原因。首先,无线业务提供多媒体通信的新业务是以数据包的形式。蜂窝系统为了和因特网连接,也需要把基于分组交换的通信引入无线移动世界中。第二,多媒体通信产生了一种全新的传输需求,而分组交换被认为是能够最有效地满足这种需求的方式。数据通信通常是突发式的,传输数据包的时间比产生数据包的时间短得多;也许需要频繁传输少量业务通信;最后,也许偶尔需要传输大文件(例如,电子邮件、网上冲浪、从 Web 站点下载文件等因特网业务)。这些传输方式的变化和基于语音的电路交换通信有很大不同。分组传输能够比电路传输更有效地应付突发业务,因为从各个用户发送过来的数据包会共用同一个传输设施,面电路交换则要求每个用户有自己独立的传输设施。但是为了保证有效地利用频谱带宽、时隙或者 CDMA 系统中的码字等无线资源,传输的建立时间必须很快,接入时间必须很短。连接时间必须比基于电路交换的 2G 蜂窝系统短。

分组交换数据通信在用户的性能需求方面也有很大不同。第 9 章讨论了电话拥塞和传递掉线率等电路交换的性能目标。分组交换有完全不一样的用户性能目标。这些目标都包含在服务质量(QoS)当中:分组优先等级(不同类型的业务可能要求不同的操作)、丢包率、包传输的延时特性以及数据的吞吐率等。在 3G 系统中,我们也将看到这些分组交换的性能需求和目标。

人们曾经认为将只有一个统一的 3G 标准来替代当前的 3 个 2G 标准。但是,全球只用一个标准很难实现,这在国际标准化进程中是司空见惯的事。现在有好几个 3G 标准同时在共存,人们期望也许将来这些标准能够统一在一起。在 10.2 节和 10.3 节中,我们将介绍 3 个 3G 标准,其中两个是基于 CDMA,另外一个是基于 TDMA。CDMA 技术有很多优点,它能动态分配带宽给不同的用户,这将会在以后讲到。这个特性在分组交换中是很有用的,能够使在支持其他业务的同时还能支持多媒体通信业务。研究表明,该特性还能在限定带宽内提供除了语音以外的更高速率的多媒体业务,这也使得蜂窝系统的性能更高。两种 CDMA 标准,WCDMA 和 cdma2000 将在 10.2 节中介绍,之前先介绍在一般情况下增加 CDMA 系统比特率的方法。cdma2000 是和 IS-95 向下兼容的。

WCDMA 源于 GSM,并被设计成用来替代 GSM 系统。许多支持 GSM 的国家和组织把通用移动电信服务(UMTS)的观念作为第三代移动通信代替 GSM 的目标。因而,缩写 UMTS 也常常等同于 WCDMA,面 WCDMA 标准也常称为 UMTS/IMT-2000 标准。这项标准由第三代合作工程(3GPP)在开发。cdma2000 标准则由之后的工程组织 3GPP2 开发。这些标准工作由欧洲、日本、韩国、美国和中国的一些标准化实体协作,分别始于 1998 年和 1999 年。

TDMA 标准又称为 2.5G/3G 标准,它能为 GSM 提供分组交换能力,从而提高射频接口的比特率,同时又尽可能地保持 GSM 系统的特征,和语音以及其他电路交换服务兼容。因此,它的变动比 WCDMA 小得多。然而,引入分组交换后,GSM 无线系统的核心网络结构需要变成以分组交换为核心的 GPRS 网络(通用分组无线业务)。GPRS 网络是靠采用一种 EDGE(全球演进式增强型数据速率)技术来提高比特率的。EDGE 技术同时被 IS-136 采用,期望通过采用该技术能够把 GSM 和 IS-136 两种 TDMA 系统合二为一。10.3 节将介绍 GPRS 和 EDGE。层结构的思想在 GPRS 的标准化中有着非常重要的地位,在其他所有的分组交换网络中也有非常重要的地位。因此,在 10.3 节中我们将简要介绍一下 GPRS 的层结构。

在 10.2 节中，我们将只讨论 3 个 CDMA 标准中的两个标准。WCDMA 分为频分复用(FDD)和时分复用(TDD)两种。频分复用在不同的频段进行上行和下行连接，频分复用在第二代移动通信中就已经被采用了。时分复用则是上行和下行在时间上交替，而占用共同的频带。在讨论 WCDMA 时，重点讨论频分复用。cdma2000 可向下兼容电路交换形式 IS-95，而且能够处理更高速率的分组交换数据。

在剩余的两章，我们将继续讨论无线系统中的分组交换。第 11 章将抛开具体的无线网络标准，讨论在普通射频链路上语音和分组数据复用的问题。这个问题通常称为复用问题，主要是资源分配问题，也就是如何有效地把频谱带宽分配给不同的用户。这通常又涉及到调度问题，在 CDMA 系统中，这关系到如何分配时间和码字；在 TDMA 系统中，每帧的时隙就是资源。为了比较不同资源的分配策略，人们做了很多工作，我们将对其中的 CDMA 和 TDMA 的部分工作做些介绍。下一节的末尾将介绍一个 cdma2000 的资源分配策略，以作为第三代移动通信中 CDMA 方案的例子。

在上行或者预留的方向中，各种复用技术很大程度上依赖于随机复用过程。随机复用在 3G 中扮演着关键角色，这将在以下章节中看到。前面讲过的，初始化一路电话也使用了随机复用。Aloha 时隙是一种最常见的随机复用形式，我们将在第 11 章中简短地介绍 Aloha 时隙的概念及技术。最后，第 12 章将讨论分组交换的无线 LAN(WLAN)，特别是 IEEE 802.11 无线 LAN 标准。这项方案被广泛用于提供因特网和其他数据网络的无线连接，使用了称为载波侦听多址(CSMA)的改进随机接入过程。802.11 已在一些商业和学术机构使用推广。在这一章的最后介绍传感器类型的无线网络，例如蓝牙技术标准。

10.2　3G CDMA 蜂窝标准

正如前面所介绍的，本节将讨论 3 种 3G CDMA 标准中的两种，频分复用的 WCDMA 和 cdma2000。首先，先简单评述一下高带宽(更高比特率)CDMA，然后分别讨论各种 CDMA 系统。

高带宽(更高比特率)CDMA 系统

3G 系统为了传输多媒体数据，需要处理更高的速率。CDMA 系统中，一般有 3 种方法来增加比特率：第 1 种方法是增加频带，这种使用多个窄频带的技术有时又称为多载波 CDMA (Milstein，2000)[①]，并在 cdma2000 内使用(但是，最初版本的 cdma2000 只有单载波)。第 2 种方法是增加码片速率，这样用比较高的扩频增益来传输更高比特率数据(回忆一下第 6 章，CDMA 系统的性能取决于高扩频增益，自从传输带宽和码片速率直接成比例后，高扩频增益有时就用码片速率比比特速率来衡量)。为了处理更高的比特率，一个简单的办法是保持码片速率不变，增加比特率，高比特率信号因而提高了扩频增益。这个原理称为变量扩频增益 CDMA (Chih-Lin I 和 Sabnani，1995a[②]，1995b[③])。因为信干比(SIR)和 CDMA 的性能直接与传输因子成

① Milstein, L. 2000. "Wideband code division multiple access," *IEEE Journal on Selected Areas in Communications*, 18, 8(August), 1344 – 1354.

② Chih-Lin I and K. Sabnani. 1995a. "Variable speading gain CDMA with adaptive control for true packet switching wireless network," IEEE International Conference on Communications, ICC95, Seattle, WA, June, 725 – 730.

③ Chih-Lin I and K. Sabnani. 1995b. "Variable spreading gain CDMA with adaptive control for integrated traffic in wireless networks," IEEE 45th Vehicular Technology Conference, VTC95, Chicago, IL, July, 794 – 798.

比例。也就是说，当比特率增加时，必须增加信号功率来补偿 SIR 的下降，否则系统性能将恶化。在 CDMA 系统的功率控制中，这一点并不很容易做到，因此必须有一个最大可能的比特率。这个值取决于传输数据的业务类型，因为不同的业务类型对比特错误率和 SIR 有不同的要求。WCDMA 使用可变增益传输换取比特率性能的增加。正如下一节我们将要看到，WCDMA 把码片速率从 IS-95 时的 1.2288 Mcps 增加至 3.84 Mcps，相应的传输带宽则需要增加到 5 MHz。

使用更高码片速率的一个优势是在利用瑞克接收机时多径会改善。在第 2 章和第 6 章中，我们曾证明了瑞克接收机能改进 CDMA 系统。在第 2 章中，证明了瑞克接收机方案利用了时间多样性，从多径环境中提取出相应能代表"RAKE""手指"的几条路径。相应每条路径的信号能够结合起来增强信号估计和改进信干比 SIR。为了使瑞克接收机更有效，信号的传输带宽必须比相干带宽大。这样会导致频率选择性衰减，引起多径回声效应。增加码片速率能增加信号带宽。以 WCDMA 的多径解析能力为例，码片速率是 3.84 Mcps，则码片持续时间是 0.26 μs。由这个值引起的不同路径，当相应路径差达到 78 m 时，不同的路径就能够被分辨出来。这大概是 240 ft 或者是市区环境下的一个典型块长度。瑞克接收机因而能够在市区微蜂窝中获得至少一个块长度的改进。第 6 章曾讲到，IS-95 能够解决多径回声，并且使用"RAKE"手指的方法改进信号检测的能力。但是，由于码片速率低于 WCDMA 使用的速率，多径分辨力受限于宏蜂窝环境。后面将看到，cdma2000 为了和 IS-95 兼容，保留了 IS-95 的码片速率。IS-95 和 cdma2000 3G 系统使用 1.2288 Mcps 的码片速率，持续时间是 0.81 μs，这个时间长度对应的是 244 m。在这种情况下的最短路径长是 240 m 左右，也就是说，瑞克接收机能改进这么长距离的性能[需要注意的是，在为了增加传输带宽而增加能分辨的路径数时，每径信号的能量会减小，在某种程度上会降低期望的增益(Milstein，2000)[①]]。

CDMA 中第 3 种获得高比特率的方法是并行多重编码。这项技术称为多重编码 CDMA (Chin-Lin 和 Gitlin，1995)[②]。在宽带 CDMA 中，每组连续信号位是并行码片编码，并且组里的每一位都用伪随机编码。编码传输序列随后一起用一个载波传输。串并转换有效降低了由组的大小所带来的更高比特的传输。因此，有一个在传输中使用的基本比特率和扩频增益结果常数。两个连续的位合并到一起，并行码片编码，每个都是用自己的码片序列，然后在扩频增益相同的情况下用双倍速率进行传输。3 个连续位集中进行传输将是基本比特率的 3 倍。这项技术在不降低扩频增益的情况下就可以增加比特率。不久，我们就将了解到，cdma2000 使用了这项技术。WCDMA 标准也使用了这技术来获得高的数据比特率。很明显，并行的码字之间必须相互正交，这样是为了每个比特都能被正确地检测，也可以防止接收机端的"自干扰"(Chih-Lin I 和 Gitlin，1995；WCDM，2000)。

多重编码 CDMA 和可变扩频增益 CDMA 之间该怎么取舍呢？多重编码 CDMA 的基本问题是传输时的宽包络变化比较大，这样对接收机的线性放大器要求就比较高。除此之外，多重编码 CDMA 的伪随机码片编码必须相互正交(Dahlman 和 Jamal，1996)[③]。多重编码 CDMA 需要多

---

① WCDM. 2000. *WCDMA for UMTS*, ed. Harri Holma and Antti Toskala, John Wiley & Sons.

② Chih-Lin I and R. D. Gitlin. 1995. "Multi-code CDMA wireless personal communication networks," IEEE International Conference on Communications, ICC95, Seattle, WA, June, 1060 – 1064.

③ Dahlman, E. and K. Jamal. 1996. "Wideband services in a DS-CDMA based FPLMTS system," IEEE Vehicular Technology Conference, VTC96, 1656 – 1660.

个瑞克接收机。多重编码 CDMA 和可变扩频增益 CDMA 的 SIR 性能是差不多的(Dahlman 和 Jamal,1996;Ramakrishna 和 Holtzman,1998)[①]。

Milstein 等在总的频谱带宽固定的前提下研究比较宽带 CDMA 和多载波 CDMA 的性能。在相同带宽上,只使用一个宽带 CDMA 信号或者使用多个窄带 CDMA,对比较的结果 Milstein 做了总结(2000)。如果宽带 CDMA(定义成传输带宽超过信号的相干带宽)使用瑞克接收机,则能够增加多径的分辨能力。这个结论下面还将讲到(所接收到的每径的能量减小了)。这必须是在相干联合的情况下。如果使用了非相干联合,则必然会带来非相干损耗(Milstein,2000)(WCDMA 使用相干联合)。另外,这些比较都假设信道参数是被完全准确地估计的。如果信道估计有误,则将使宽带系统性能变差。多载波系统的每个频带用一个瑞克接收机,其性能将比单载波宽带系统提高不少。虽然接收机的结构更复杂了,但是并行处理速度可以做得很低,而宽带系统串行处理的速度则非常高(Milstein,2000)。当宽带 CDMA 和多载波 CDMA 的 SIR 都大于 0 dB 时,其性能差不多。当 SIR 变差时,多载波系统性能会更好些。

WCDMA

接下来将详细讨论各种 CDMA,首先讨论频分复用的 WCDMA。和第二代移动通信 IS-95 一样,这个系统是直接序列 CDMA 系统,但是码片速率从 IS-95 的 1.2288 Mcps 上升到 3.84 Mcps,从而在扩频增益不变的情况下可以有更高的传输比特率。当然,高的码片速率要求更高的传输带宽,该系统的传输带宽是 5 MHz。因此,分配给 WCDMA 的频带值也必须提高。由于这个缘故,欧洲、日本和韩国选择了 2 GHz 的频带,然而美国并没有把这个频带分配给 3G 系统。1920 MHz ~ 1980 MHz 这 60 MHz 被选择用于移动台至基站的上行传输,2110 MHz ~ 2170 MHz 这 60 MHz 用于基站至移动台的下行传输。WCDMA 的设计意图是尽管使用 CDMA 技术和更高的频带,但应尽可能多地和 GSM 保持一致。WCDMA 系统因此还设计出从 WCDMA 网络到 GSM 网络的切换。和 IS-95 一样,WCDMA 也支持软切换(注意到,单词 handover 在北美之外被 handoff 替代,在更多的国际文件中使用的是 handover 而不是 handoff 这个术语)。

WCDMA 在广域环境下支持 384 Kbps 传输速率,在局域环境中支持 2 Mbps。注意,这个速率比在 2G 系统中用电路交换的语音速率高得多。但是,由于包头的开销、包队列延时、包长度的可变性会导致没有充分利用帧结构和时隙结构,实际上的传输速率可能达不到这个速率。随后,在讨论 IS-95B(cdma2000 家族的第 1 个成员),GPRS,EDGE 技术时,我们将对这点做简要的量化分析。

WCDMA 是如何做到这样高的速率的呢? WCDMA 码片速率固定在 3.84 Mcps,使用可变扩频增益和多重编码技术。可变扩频增益技术的传输数据速率有一个可变范围,上行可以从 15 Kbps 调节到 960 Kbps,下行速率最大可以达到 1.92 Mbps。下行速率比较高的缘故是在这个方向上使用了 QPSK 调制,上行则使用的是 PSK。所有的传输数据使用 1/2 码率的卷积码来保证用户数据的优先权。实际用户数据比上面讲到的一半还要低,例如,如果包括前向纠错编码,下行速率的范围就比从 7.5 Kbps 到 480 Kbps 低。扩频增益定义成码片速率除以传输比特率(WCDM,2000),也从 256(3.84 Mcps/15 Kbps)降低至 4(3.84 Mcps/960 Kbps)。如果需要,6 个

---

① Ramakrishna, S. and J. M. Holtzman. 1998. "A comparison between single code and multiple code transmission schemes in a CDMA system," VTC98, May, 791 – 795.

码字可以并行使用,以获得更高的比特率(这就是前面所说的多重编码)。扩频增益为4的上行信道的最大可能比特率是 6×960=5.76 Mbps。用户的信息速率稍微低于这个值的一半(这就是前面所讲的多重编码过程)。由于扩频增益比较低,这么高的传输速率只能在局域环境中。广域环境的最大传输速率在 384 Kbps 左右。

前面已经讲过 3G 系统在设计的时候考虑了服务质量(QoS),同时也已经讲过 WCDMA 通常意义上是和通用移动电信业务(UMTS)相同的。WCDMA 处理4类 QoS 业务,其中两类业务是实时的,另两类是非实时的(WCDM,2000)。实时业务包括传统的那些业务(通常这些业务是双向交互的),如语音与视频业务和多媒体流业务。这些传统业务的端到端传输延时很短,组成业务流的连续信息实体随时间的变化不大(WCDM,2000)。通常认为实时语音业务的端到端传输延时不能大于 100 ms,当大于这个延时时,人们就能感觉出通信出现了问题。流业务对连续信息实体之间的时间要求也很严格(WCDM,2000)。

非实时业务包括交互业务例如 web 浏览,后台业务例如电子邮件(WCDM,2000)。所有这些业务对数据完整性要求比较高,也就是说,接收到信息的错误率必须很低(相反,即使5%的语音数据在传输中出现错误,语音仍然能被正常识别)。和交互业务不同的后台业务(如电子邮件)对传输延时并不敏感。在因特网业务中,这种业务又称为"尽力而为"业务。非实时交互业务通常具有"突发"的特性。在相当长的时间内,也许只有一小段时间需要传输数据,而别的时间都是空闲的。例如,移动客户要想连接因特网服务器(例如,Web 地址)需要向服务器发送一个短的请求消息,接着等待服务器的答复,接收到答复后再发送一个请求。我们将看到在 WCDMA 系统中,在发送小数量数据包的时候也采用了这种规定。

正如上面提到的 WCDMA 上行链路(移动台到基站)使用 PSK 调制,下行链路使用 QPSK 调制。也就是说,下行链路的传输速率是上行链路的两倍,在因特网、Web 地址业务中,同样具有类似的特点。终端采用相关检测,通过导频信息做信道估计,为瑞克接收机提供所需的信道信息。

在无线信道中,使用重复的帧结构来传输数据。数据帧是 10 ms 长,每帧的具体比特率可能会有所不同。但是,2G 系统中(第8章)语音编码(编码/解码)的帧长是 20 ms。可以用8种不同比特率中的一种来传输语音信号,范围是 12.2 Kbps 到 4.75 Kbps。速率也许是每隔 20 ms 帧就变化一次。选择 12.2 Kbps 是为了和 GSM 系统兼容,选择 7.4 Kbps 是为了和 IS-136 兼容;而 6.7 Kbps 是和日本的 PDC 语音码率一样的(WCDM,2000)。WCDMA 使用代数编码兴奋线性预测技术(ACELP),处理 8000 语音样本/s 或者 160 样本/20 ms 帧(WCDM,2000)。

现在重点讨论分组交换数据业务的可变速率。如果该业务通过无线信道传输,其结果会怎样呢?为此,使用了两种专用信道(DCH)类型,它们都是专用的物理信道。一种 DCH 类型用于传输用户数据和高层控制通信,但是这种高层控制通信有别于通过无线信道物理操作的控制通信(分组交换的层结构概念将在 10.3 节与 GPRS 结合起来讲)。这个逻辑信道所映射的物理信道是专用物理数据信道或者 DPDCH,在给定的频率用具体的编码,并且为单用户保留。DPDCH 使用可变传播因素技术和多重编码技术(必要时),其比特率有可能会在一个 10 ms 帧到下一个 10 ms 帧时变化。因而,这种信道支持4种不同的业务类型,可根据需要提供带宽。

第2种 DCH 是在一个蜂窝小区内所有的用户或者某一部分特定的用户使用共同的逻辑信道。为了支持可变速率数据传输 DCH,需要提供诸如 DCH 正在使用的数据传输率等信令和控制信号。这种信令/控制 DCH 将映射到 DPCCH 中。DPCCH 也包括给定频率的具体编码,但是通常运行在扩频增益因子为 256 的固定比特率下。它是和 DPDCH 并行传输的。DPCCH 携

带着同一帧时刻 DPDCH 的传输比特率的相关信息。DPCCH 的上行和下行的结构是不同的。图 10.1 描述了 WCDMA 10 ms 重复帧的上行 DPCCH 结构(WCDM,2000)(携带用户数据或者上层控制信号的 DPDCH 也画在图中,与 DPCCH 平行)。注意,每帧被分成 15 个时隙,每个时隙都是 2/3 ms 宽,由 2560 个码片组成。

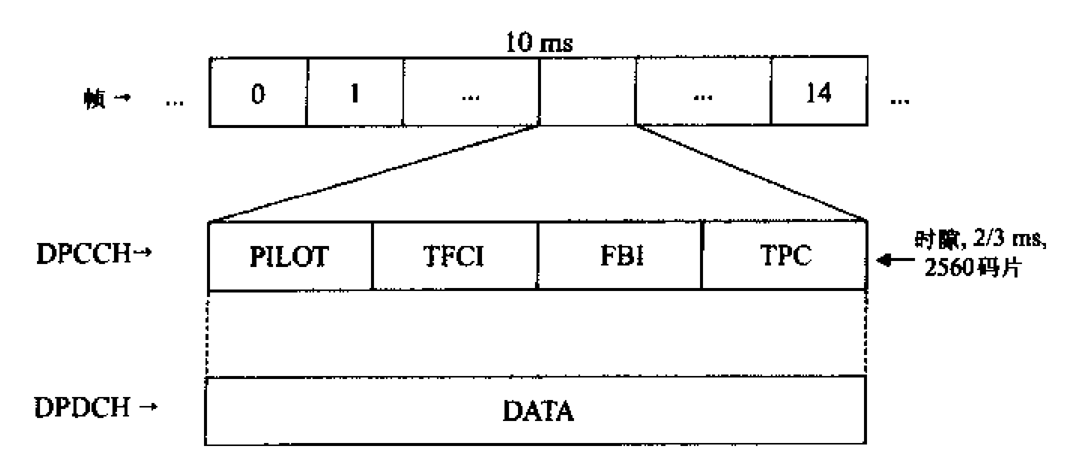


图 10.1 WCDMA 上行专用物理信道(引自 WCDM,2000:Fig.6.11)

每个 DPCCH 隙由 4 个区域组成,其中 FBI 域是可选的。导频域用来在相干检测和 RAKE 接收中做估计无线信道用。传输格式联合指示器(TFCI)域指示伴随 DPDCH 帧的数据传输率。传输功率控制域(TPC)提供下行功率所需的控制信息。如果想用多天线来改进分集阵列,就用可选反馈信息(FBI)域来控制天线的相位和幅度。传输 DPDCH 帧时,每 10 ms 帧就用比特交织来防止衰减。如果允许更多的传输延时,则可以在 20 ms、40 ms 或者 80 ms 帧处也插入比特交织(传统业务的发送和接收都对端到端的传输延时有严格限制,但是非实时业务的端到端的传输延时限制则要松得多)。因为 DPCCH 信道同时携带有 DPDCH 率信息,所以当 DPCCH 域第 1 次被解开后必须把 DPDCH 比特保存在接收的基站内,详细内容请参见 WCDM(2000)。下行帧结构不同于上行帧结构。所有的 DPCCH 和 DPDCH 以 15 时隙、10 ms 帧在下行物理信道中时间复用,并且没有使用 FBI 域。每个时隙用 TFCI,TPC 和导频域来替代 DPDCH、数据域。图 10.2 说明了这种下行专用的物理结构。同样,我也可以参考 WCDM(2000)获取更多的信息,包括上行下行的帧格式,以及它们各自的接收机是如何工作的。

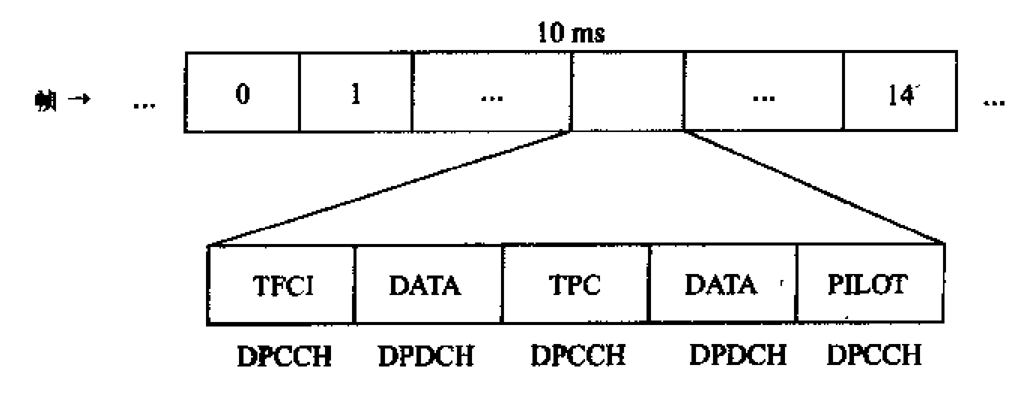


图 10.2 WCDMA 下行专用物理信道:控制和数据信道时间复用(引自 WCDM,2000:Fig.6.14)

至此,在讨论 WCDMA 中使用了专门的物理信道来进行分组数据传输。在 2G 系统中,控制信道也用来传输接收机与基站同步所需的物理层信息、获取码字信息、寻呼、用随机复用建

立电话的初始连接,等等。这些都是能被所有接收机利用的通用信道,而不是前面讨论的那种专门用来传递数据的信道。为此,定义了6个逻辑信道,其中两个是可选的,除此之外还有一系列纯物理信道(WCDM,2000)(回忆在第8章讲到的IS-95,物理信道定义为在具体频率下的特定的码字)。表10.1列出了这些信道,同时标明方向的选择性:上行(U)、下行(D)。注意,只有上行信道是随机复用的,也就是说,随机复用只使用在移动台向基站建立连接时。更多信息请看WCDM(2000)。

表10.1 WCDMA通用信令/控制信道

逻辑信道	指示	物理信道
BCH	D	PCCPCH,基本公用控制物理信道
FACH*	D	SCCPCH,第二公用控制物理信道
PCH*	D	
RACH	U	PRACH,物理随机接入信道
可选		
DSCH	D	PDSCH,物理下行共享信道
CPCH	U	PCPCH,物理公用分组信道
	D	SCH,同步信道
	D	CPICH,公用导频信道
	D	AICH,捕获指示信道
	D	PICH,寻呼指示信道
	D	与CPCH过程相关的4个附加信道

下行广播信道BCH携带随机复用码字并在一个蜂窝小区内可接入时隙,这也是移动终端所需要的。不管终端性能如何,所有终端都需要接收这路信道并且译出它所携带的信息。除此之外,这个信道必须能够连接这个蜂窝小区内所有的用户。因此,这个信道通常用低速、固定数据速率、相对比较高的功率传输。前向接入信道(FACH)一旦完成随机接入过程就传输信令和控制信息。有些分组数据也会借助这个信道传输,例如突发数据就会通过这个信道从基站传输到移动台。一个蜂窝内也许有很多信道,但是至少需要有一个信道的速率是很低的,使得所有移动台都能够接收到这个信道。别的FACH传输速率也许会高一些。快速功率控制在这个信道中没有使用。注意表10.1中FACH和下行寻呼信道(PCH)共享同一个物理信道。两个信道可以在一个物理信道中时间复用,或者FACH也可以有它自己独立的物理信道。

PCH和在前面所讨论的2G系统中的寻呼信道有着相似的功能。寻呼信道是和寻呼指示信道(PICH)联合使用的,能够使接收终端从"睡眠"模式转化到保持功率。接收终端周期性地唤醒并且监听PICH,以决定是否有寻呼消息在等待。细节请看WCDM(2000)。

2G中的上行随机接入信道(RACH)是移动台用来初始化一路电话,告知移动台目前在一个新的区域,或者当开机时注册网络。捕获指示信道(AICH)是基站用来回答RACH消息的。当出现RACH消息时,它回射一个签名序列(一个AICH应答能回复16个不同的签名序列)。公用导频信道(CPICH)从基站到移动台下行连接,能够给移动台检测AICH提供相位参考和信道附加信息。RACH消息有10 ms或者20 ms长,能够传输一个简单的大概是600比特长的突发数据包;传输因子从256下至32,也许能够用在随机接入信道中。若传输因子是32,则对应的传输比特率是120 Kbps。如果考虑1/2码率卷积码,则10 ms能传输1200比特。如果考虑功

率因素,使传输因子达到128,则在RACH消息内要传递的信息就可以减少到150比特。

同步信道SCH用来蜂窝搜索、获取必要的帧和时隙同步。实际上SCH使用了两个同步信道,主信道使用了在一个系统中所有蜂窝内都相同的256个码字传输序列,而辅信道各自使用了不同的码字,两个信道是并行传输的。SCH和主公用控制信道(PCCCH)时间复用,PCCCH用来传输BCH(见表10.1)。它使用10 ms帧内的1/10时隙,也就是一个时隙内2560个码字中的256个。

注意,表10.1列出的可选公用分组信道(CPCH)既处理上行信道又处理下行信道。这个方法使用的是改进过的带碰撞检测能力的随机接入技术,该技术优于通常使用的RACH方法,能传输更多的数据。我们将在讨论时隙Aloha的随机接入方法时,具体分析这种技术的优点。可选的下行共享信道(DSCH)为专门用户提供数据下行业务,在传输高速率、高突发业务时特别有用。正如前面讲到的,这种数据是偶然传输的分组形式的消息,只占时间的一小部分,例如Web服务器响应客户端的接入请求、文件传输。这个信道通常是和下行专用信道(DCH)相联系的,多个用户共享该信道。和DCH差不多,它提供可变比特率、支持快速功率控制。

接入方法总结如下:移动终端为了接入移动网络,对BCH进行解码,决定可以使用的RACH,同时收集码字并设置签名序列,RACH和签名方式都是随机的。终端同时测量下行链路的功率等级,用这个测量值来初始化RACH传输功率等级。然后,传输1 ms RACH前导码,等待基站的AICH响应。如果RACH前导码被基站正确接收,AICH将使用和前导码一样的签名方式。移动台随后会发送完整的RACH消息,消息是10 ms或者20 ms长。如果在几个时隙之后,移动台不再接收到AICH消息,它将重新传输1 ms带有同样签名序列的前导码,但是使用更高的传输功率,重复这个接入过程(连续功率增加的数是1 dB的整数倍,由基站的原始BCH信息指定)。

cdma2000

cdma2000系统设计成向下兼容IS-95系统(第二代CDMA系统,在前面的章节已经讨论过)。cdma2000系统的发展和标准化按时间也许可划分成3个阶段,尽管这些系统按一般的设计准则都属于cdma2000家族,但是不同系统的设计方法不同,引导系统朝更高的数据速率发展。我们将在下面依次对各个系统进行讨论,并对各系统的分组数据能力做着重分析。所有的这些系统都使用前面10.1节所描述的cdma多重编码的方法来提高数据率,同样采用了可变扩频增益码和多载波方法。

IS-95的第1次扩充使用高比特率分组数据业务,称为IS-95B。该系统的前向或者下行(基站到移动台)分组数据业务能达到115.2 Kbps。严格意义上说,cdma2000是提供更高速率的IS-95。cdma2000之后分成两种空中接口模式,cdma2000 1X和cdma2000 3X。cdma2000 1X使用和IS-95、IS-95B相同的片上码字速率1.2288 Mcps和1.25 MHz带宽。cdma2000 3X是cdma2000家族里的多载波版本,前向使用了3个1.25 MHz载波,覆盖5 MHz带宽来提高传输率。每个载波使用IS-95的片上速率1.2288 Mcps。上行方向则只使用一个载波,片上速率是3.6864 Mcps(3×1.2288 Mcps)或者1.2288 Mcps。cdma2000家族的第3个成员1×EV-DV是一个专门用来分组数据通信的高速率系统。

cdma2000使用前向和逆向分别等价于下行和上行。在下面,我们也将用前向逆向来替代上行下行。在讨论cdma2000业务和信令信道时,用F来代表前向信道,用R代表逆向信道。

从 IS-95B(Sarikaya,2000①;Tiedmann,2001)②开始,这是 IS-95 的分组模式,也是 cdma2000 的先驱。它的前向和逆向传输速率都是 115.2 Kbps,带宽和片上速率与 IS-95 相同,分别是 1.25MHz 和 1.2288 Mcps。这个系统支持 64 个物理信道,每个物理信道有两个连接方向。这些信道有两种类型,公用信道和专用信道,和 WCDMA 一样(请读者自己比较一下第 8 章描述的 IS-95 和 WCDMA 的信道)。

首先考虑公用信道,这些和第 8 章所介绍的 IS-95 下的信道是一样的。在前向信道中,这些信道包括前向导频信道(F-PICH)、前向同步信道(F-SYNCH),并且增加到 7 个前向寻呼信道(F-PCH)。逆向公用信道组成随机接入信道(RACH),它和 IS-95 中的一样用来作为寻呼信道。F-PICH 用来提供信道估计信息,包括功率控制测量以及其他信道追踪和移交所需的信息。F-SYNCH 是 1.2 Kbps 信道,提供系统时间和寻呼信道数据率等参数。RACH 仍用来初始化一路电话、响应寻呼消息以及在本地注册,这些操作的速率都是 4.8 Kbps。

接着考虑专用业务信道,它分两种类型:基本信道(FCH)和补码信道(SCCH)。SCCH 使用多重编码技术来增加比特率。FCH 的数目随着数据的传输而变化,取决于 64 个信道中有多少用于传输信令和控制。前向信道中必须产生一个 F-PICH。如果 F-SCH 和 F-PCH 也产生了,则前向基本信道 F-FCH 将会减少。逆向信道的 FCH 是 64 个,少于 RACH 的数目,每个对应一个 PCH。

FCH 在两个方向上的传输速率最高是 9.6 Kbps 或者 14.4 Kbps(注意,14.4 Kbps 比 IS-95 的传输速率要高)。SCCH 的传输速率通常是其对应的 FCH 的最大速率,特别地,如果 7 个 SCCH 对应一个 FCH,并且 FCH 的最大传输率是 14.4 Kbps,那么传输率就是 $8 \times 14.4 = 115.2$ Kbps。注意多重编码是如何增加比特率的。

移动台使用逆向信道来初始化分组数据传输。如果在逆向信道中移动台需要更高的传输率,则它在 R-FCH(Sarikaya,2000)信道内向基站发送一个增补信道请求消息。基站应答这个消息,分配一些反向补码信道(R-SCCH)给移动台。另一方面,如果基站决定它需要 F-SCCH 来增加移动台的速率,则它会发送补充分配消息给移动台,开始使用分配的 F-FCH 和 F-SCCH 来传输。

然而,实际有效的数据吞吐率低于实际传输的值,因为数据包的头是额外的开销;分组数据并不完全适合于 IS-95 时隙(Sarikaya,2000)。实际数据吞吐率低于理想值也是所有分组数据交换系统网络的特征。这点已经在先前的 WCDMA 比特率中提到过,并且在下面的 GPRS 和 EDGE 中也会碰到。

例子

用一个简单的例子来说清这一点。IS-95 中帧或者时隙是 20 ms 长,如果速率是 14.4 Kbps,那么每帧长度就是 288 比特。假设要传输一个 1200 字节或 9600 比特的分组数据,这个分组数据必须分成更小的段来填充到连续的帧里面。在段内,每帧的大小不能超过 32 个字节,31 个字节用来装数据,剩下的一个字节是标明该段在分组数据中的位置。为了达到 115.2 Kbps 的传输率,用 7 个 SCCH 和一个 FCH,那么需要 5 个 20 ms 帧来传输这个数据包,需要的时间是 0.1 s,有效传输率是 96 Kbps。分组越长,有效传输率也就越高;短的分组会降低有效传输率。例如,请读者自己计算在同样条件下 800 字节数据的有效传输率是 80 Kbps。如果考虑反向信道的随机接入过程,则从移动台到基站的传输时间会更多,有效传输率会更低。

① Sarikaya, B. 2000. "Packet mode in wireless networks: overview of transition to third generation," *IEEE Communications Magazine*, 38, 9(September), 164 – 172.

② Tiedmann, E. G., Jr. 2001. "cdma20001X: new capabilities for CDMA networks," *IEEE Vehicular Technology Society News*, 48, 4(November), 4 – 12.

接下来转到 cdma2000 中的 3G 版本(cdma,2002)。这种系统分成 2 种空中接口模式,1X 使用和 IS-95 以及 IS-95B 一样的扩频率(1.2288 Mcps)和带宽(1.25 MHz);而 3X 使用多载波。只讨论 cdma2000 1X,因为 1×EV-DV 的进展已经使 3X 不再那么被需要了(Tiedmann,2001)。这个系统为了和 IS-95 及 IS-95B 兼容,和它们保持了很多一致性。例如,调制以及码的类型和 IS-95 很相似。IS-95B 专门的信道和公用的信道仍继续被使用,为了增加比特率同时改进系统性能又增加了几个信道。尤其是引入了增补信道(SCH),SCH 使用扩频增益,使每个信道能够达到 307.2 Kbps。通常一个 FCH 配两个 SCH。为了对抗阴影衰落,通过提高功率控制,前向链路连接仍可以达到 800 bps 的速率,同时能够通过改进传输分集能力来提高空时扩频性能。一般通过在两根正交天线上重复调制符号,不需要额外的 Walsh 函数。在低速移动终端上,采用这种技术能够大幅提高性能。在发射端使用 1/2 或者 1/4 码率的卷积码,以及将 turbo 码用于高速率的 SCH,使每帧携带超过 360 比特的数据,都能够提高性能(cdma 2002[①];Tiedmann,2001;Willenegger,2000[②])。在第 7 章中我们讲到,使用 turbo 码能相当可观地提高系统性能,1/4 码率的卷积码也能改进系统性能。cdma2000 中最常用的帧长是 20 ms,这和 IS-95 一样。帧长为 5 ms,40 ms 和 80 ms 的也能被使用。

cdma2000 中定义了各种无线配置(cdma,2002),每种配置都详细规定了业务信道所能使用的数据率范围、纠错技术(卷积码或者 turbo 码)以及调制技术。首先讨论反向信道,这个方向有 6 个具体的无线配置,4 个是 1X 系统,2 个是 3X 系统。配置 1 与带 FCH 和 FCH 相对应的 SCCH 信道的 IS-95B 系统相对应,更高比特率的配置 3 支持额外的 SCH 信道。配置 1 能够支持配置 2;配置 3 能支持配置 4。表 10.2 列出了 cdma20001X 中 4 种反向射频信道的具体配置(cdma,2002)。注意配置 1 为每个 FCH 提供了 9600 bps 的最大速率,其相应的 7 个 SCCH 也均能达到 9600 bps。配置 2 为每个 FCH 提供了最高 14.4 Kbps 的最大速率。配置 3 中每个 FCH 有两个 SCH,每个 SCH 支持的最高速率为 307.2 Kbps,FCH 的最高速率为 9600 bps。另外,有 5 个可能的 SCH 速率能达到 9600 bps 以上,而且这些数据率均是 9600 bps 的整数倍,等于 $9600\times N$, $N=2,4,8,16,32$;分别是 19.2 Kbps,38.4 Kbps,76.8 Kbps,153.6 Kbps,307.2 Kbps。用可变扩频增益来产生更高的数据速率,当每比特的码片速率降低了,码片速率仍然固定在 1.2288 Mcps。特别地,如表 10.2 所示,扩频增益以 $128/N$ 的方式下降,当 $N=32$ 时降至最小值 4。配置 4 和配置 3 很相似,也同样支持 SCH 信道,但是速率更高,其速率是 14.4 Kbps 的整数倍,最高可以到 $16\times 14.4=230.4$ Kbps。在这个速率下相应的扩频增益是 $85.33/16=5.3$。在配置 3 和配置 4 中得到的最高速率信道可以用 turbo 码替代卷积码。

在图 10.3 中画出了配置 3 中 5 个最高速率 SCH 的调制符号的产生过程(cdma,2002),并且假设帧长是 20 ms,这副图是和图 8.19 做对比的。图 8.19 处理 IS-95 的前向业务信道,其速率要低得多。注意每 20 ms 帧间隔的信道比特数的变化取决于所需要的数据比特率。16 个循环冗余校验位(CRC)用来衡量帧质量,加入到输入信道比特位中,结果数据块反馈到卷积或者 turbo 码中,用来纠错。每帧的末尾还加了 8 个比特。表 10.2 的数据率作为卷积码和 turbo 码的结果输入比特率。每帧间隔内的数据率也许都会变化。cdma2000 中编码器的长度一般限于

① cdma 2002. Physical layer standard for cdma2000 spread spectrum systems, Release A, 3GPP2 C.S0002-A, Version 6.0, 3rd Generation Partnership Project 2, 3GPP2, February, http://www.3gpp2.org/public_html/specs/

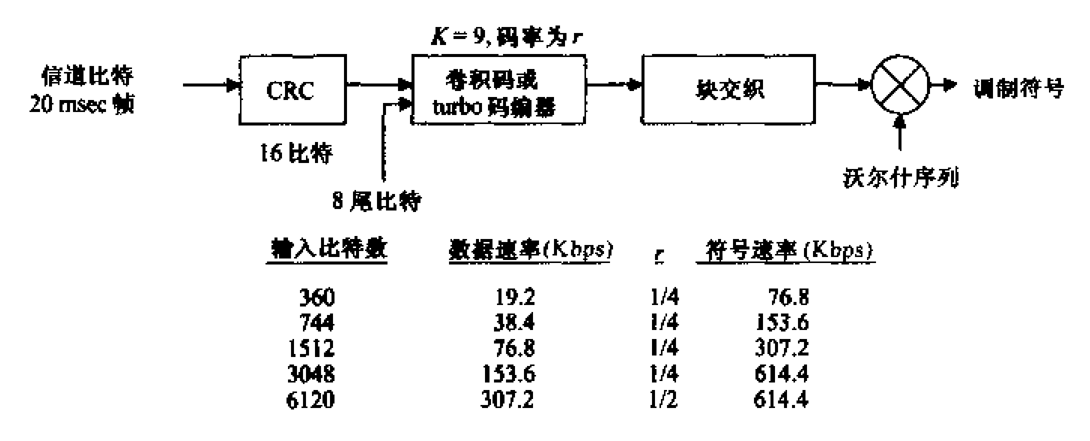
② Willenegger, S. 2000. "cdma2000 physical layer: an overview," *Journal of Communications and Networks*, 2, 1(March), 5-17.

$K=9$,其 r 的变化取决于无线配置和数据率,参见表 10.2。特别地,在 RC3 的 SCH 中 $r=1/4$,除了最高的数据率 $r=1/2$。数据率也许每帧间隔都会变化。

表 10.2 反向射频信道配置(RC),cdma2000 1X

RC	数据速率(Kbps)	卷积码/turbo 编码器(速率为 r)	扩频增益
1	1.2, 2.4, 4.8, 9.6	卷积码, $r=1/3$	128
2	1.8, 3.6, 7.2, 14.4	卷积码, $r=1/2$	85.33
3	1.2, 1.35, 1.5, 2.4, 2.7, 4.8, 9.6	卷积码, $r=1/4$	
	$9.6\times N$, $N=2, 4, 8, 16$ (9.6, 19.2, 38.4, 76.8, 153.6)	卷积码/turbo, $r=1/4$	$128/N$
	307.2 (9.6×32)	卷积码/turbo, $r=1/2$	4
4	1.8, 3.6, 7.2, 14.4	卷积码, $r=1/4$	
	$14.4\times N$, $N=1, 2, 4, 8, 16$ (14.4, 28.8, 57.6, 115.2, 230.4)	卷积码/turbo, $r=1/4$	$85.33/N$

编码输出比特取决于使用的数据率,并且和可变长度块相结合,块间隔见图 10.3,它们使用 Walsh 编码。如果使用 R-FCH,那么在 Walsh 编码之后,图 10.3 中 SCH 输出的调制符号会加到 FCH 调制符号中,同时还传输反向导频信道(R-PICH)信号。正交 Walsh 编码用在反向增补信道(R-SCH 和 R-FCH 符号)中,以保证它们在基站接收系统中能够被分离。总的信号然后用 1.2288 Mcps 的伪随机码片序列扩频。每个这样的序列都是由各自移动台的长码字发生器惟一产生的,这和第 8 章中 IS-95 所描述的比较相似。R-PICH 是一个连续传输、引入了直流的信号,能够在接收基站相干解调,因而能够改进反向信道的检测能力。



输入比特数	数据速率(Kbps)	r	符号速率(Kbps)
360	19.2	1/4	76.8
744	38.4	1/4	153.6
1512	76.8	1/4	307.2
3048	153.6	1/4	614.4
6120	307.2	1/2	614.4

图 10.3 反向增补信道结构,无线配置 3,cdma2000(引自 cdma,2000:Fig.2.1.3.1.1.1-8)

其他 cdma2000 反向信道包括:反向专用控制信道(R-DCCH),使用它在配置 3 和配置 4 中传输 9600 bps 或者 14 400 bps 的突发数据或者信令信息;反向接入信道(R-ACH),它与 IS-95 和 IS-95B 中所使用的比较相似;反向增强接入信道(R-EACH),以及反向公用控制信道(R-CCCH)。除了 R-ACH 信道之外,别的信道都是和 R-PICH 一起传输的。R-EACH 支持两种接入模式:基本接入模式与 ACH 比较相似,但是允许短的消息作为随机接入过程的一部分被传输;保留接入模式用来加速接入过程,并且允许更长的消息作为接入过程的一部分被传输(Tiedmann,2001;Willenegger,2000)。在预留接入模式下,移动台首先使用基本接入模式向基站

请求分配容量。基站使用前向公用分配信道(F-CACH)回答一条分配消息。F-CACH 信道用来加速接入确认和更快地解决接入碰撞。一旦终端接收到这个分配消息,它就使用 R-CCCH 信道传输所期望的数据。移动终端功率是由基站控制的,基站通过前向公用功率控制信道(F-PCCH)的子信道的传输进行控制,而 F-PCCH 携带有功率控制子信道,每种都和 R-CCCH 惟一对应(Willenegger,2000)。这种功率控制方法作为预留接入模式的一种用来降低接入过程中的干扰,提高信道的可靠性。图 10.4 总结了 cdma2000 1X 中所使用的预留信道(cdma,2002)。各种信道的块图结构和第 8 章中的 IS-95 非常相似,反向增补信道(R-SCH)的扩展图(见图 10.3),具体情况可参阅 cdma(2002)。

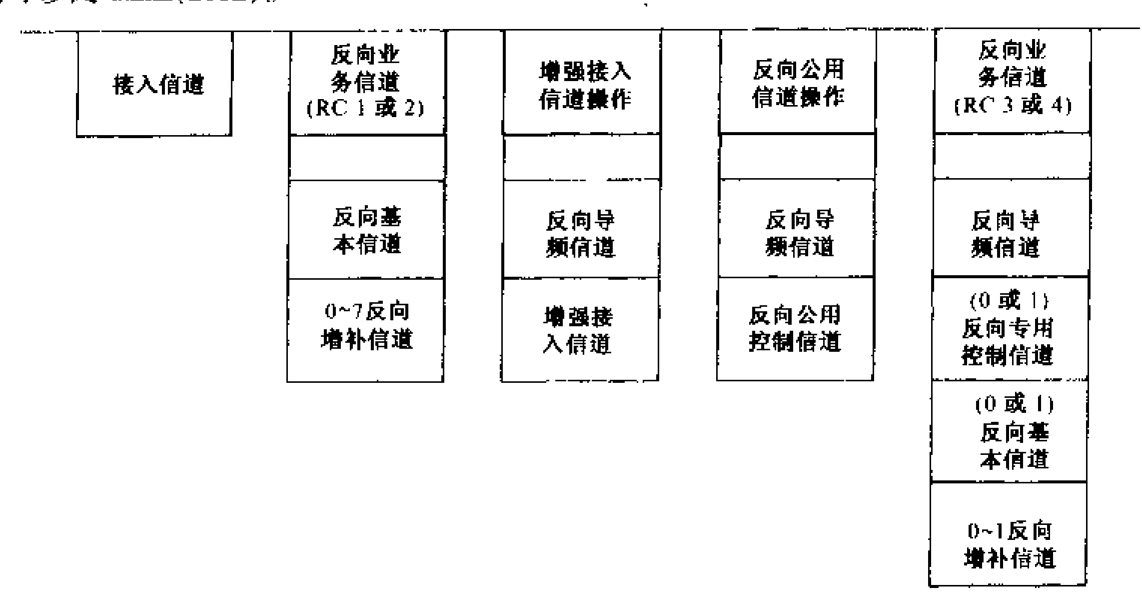


图 10.4 反向信道,cdma2000(引自 cdma,2002:Fig.2.1.3.1.1-1)

现在考虑 cdma2000 1X 前向信道。业务信道继续被使用,但是它们的特征被无线配置(cdma 2002;Willenegger,2000)重新定义了。现在 cdma2000 1X 有 5 个无线配置来替代反向的 4 个配置,参见表 10.2。表 10.3 同时显示了用于反向的相应的 RC,因此配置 1 和配置 2 分别和反向的 RC1 和 RC2 相联系。这些又代表委派配置,向后与 IS-95 和 IS-95B 兼容。前向 RC3 和 RC4 都与 RC3 以及反向 RC3 相联系;前向 RC5 与反向 RC4 相联系。RC1 和 RC2 的业务信道包括前向基本信道(F-FCH)、前向补码信道(F-SCCH),每个前向业务信道最多对应 7 个 F-SCCH。这些信道具有在反向业务信道中同样的特征,并且仍然向后兼容 IS-95 和 IS-95B。高比特率前向无线配置 RC3 到 RC5 的业务信道与在反向信道中的非常相似:每个前向业务信道会携带一个低比特率的 F-FCH 和两个以下的高比特率前向增补信道。SCH 的数据率会随着帧而变化,这与在反向业务信道中一样。如表 10.3 所示,我们再次使用可变扩频增益技术来产生可变比特率,扩频码片速率仍然固定在 1.2288 Mcps。前向增补信道(F-SCH)运行模式与在反向增补信道(R-CH)中的非常相似,参见图 10.3。图 10.3 描述了前向业务信道的产生,以及 cdma 中的前向信令和控制信道(2002)。

前向控制和信令信道构成了前向控制信道,这和前面第 8 章所讨论的 IS-95 类似,但是又加入了像 F-PCCH 和 F-CACH 等新的信道。回忆一下,第 8 章中 IS-95 除了有前向业务信道,还有前向的导频、同步和寻呼信道。这些信道也都在 cdma2000 结构中出现。新定义的信道除了 F-PCCH 和 F-CACH 以外,还包括一个前向公用控制信道(F-CCCH)、一个前向广播信道

(F-BCCH),F-BCCH 用来传输寻呼消息、分配信道、短广播消息以及消息头。寻呼过程分为两步,这样可以降低移动功耗、增加电池寿命。移动台通常是在睡眠模式,一个快速寻呼信道(F-QPCH)由连续单比特指示器组成,用来通知寻呼消息的到达。移动台被唤醒,检查指示器,这只需要用很少的功率。移动台随后读取寻呼信道的消息,在向后和 IS-95 兼容的系统中读取 IS-95 类型的寻呼信道,在更新的 cdma2000 系统中读取 F-CCCH 信道。其他的公用信道包括前向发送分集导频信道(F-TDPICH)、前向辅助导频信道(F-APICH),以及前向辅助发送分集导频信道(F-ATDIPCH),这些信道用来使基站能够配备点波束和具有空间分集技术的多重天线(cdma,2002;Tiedmann,2001;Willenegger,2000)。

表 10.3 前向无线配置(RC),cdma2000,1X

RC	数据速率 (Kbps)	反向信道 RC	卷积码 /turbo 编码器(速率为 r)
1	1.2, 2.4, 4.8, 9.6	1	卷积码., $r = 1/2$
2	1.8, 3.6, 7.2, 14.4	2	卷积码., $r = 1/2$
3	1.2, 1.35, 1.5, 2.4	3	卷积码., $r = 1/4$
	2.7, 4.8, 9.6		
	19.2, 38.4, 76.8, 153.6		卷积码./turbo, $r = 1/4$
4	1.2. 1.35, 1.5, 2.4	3	卷积码., $r = 1/2$
	2.7, 4.8, 9.6		
	19.2, 38.4, 76.8, 153.6		卷积码./turbo, $r = 1/2$
	307.2		
5	1.8, 3.6, 7.2, 14.4	4	卷积码., $r = 1/4$
	28.8, 57.6, 115.2, 230.4		卷积码./turbo, $r = 1/4$

cdma2000 同时设计了新的前向专用信道,包括专用控制信道(F-DCCH)以及前向专用辅助导频信道(F-DAPICH)。F-DCCH 用来传输突发数据或者信号信息,F-DAPICH 的作用和公用控制信道中的 F-APICH 类似(Willenegger,2000)。所有的前向信道都列在图 10.5 中(cdma,2002),图 10.5 同时标明了每个基站所拥有的信道数。

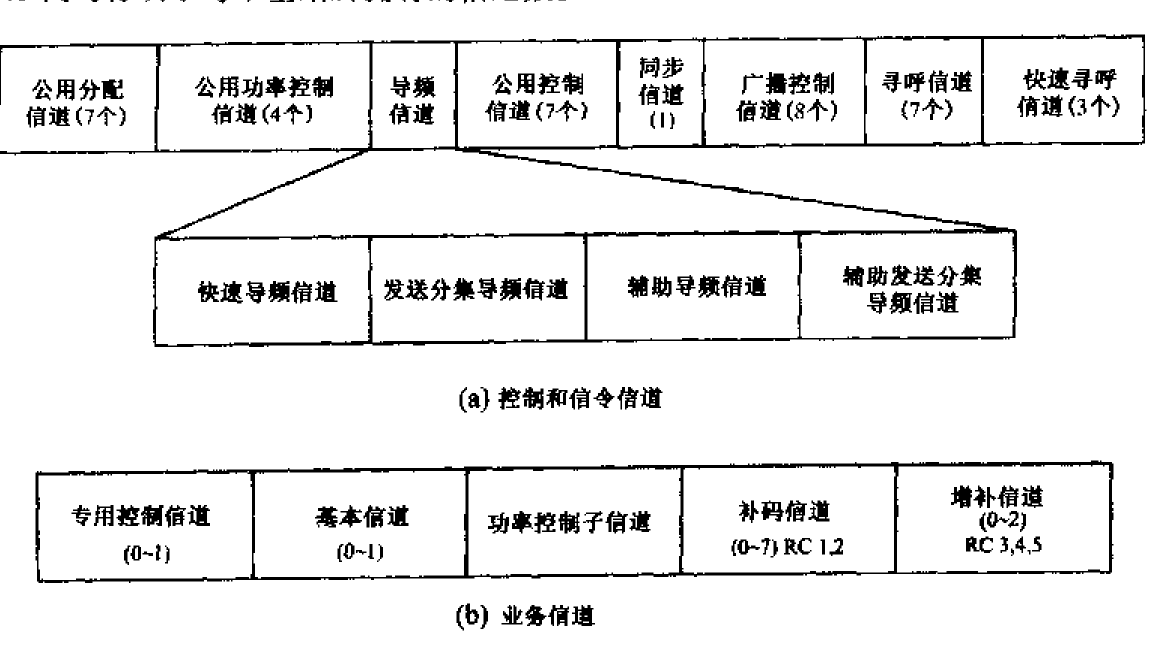


图 10.5 前向信道,cdma2000(引自 cdma,2002:Fig.3.1.3.1.1 – 1)

移动台在刚开机以及注册时获取系统信息，相应的信道是如何工作的呢？基站连续传输常数 1 和 PN 序列，代表传输前向导频信道(F-PICH)。和 IS-95 中一样，不同的基站由不同的 PN 偏移来识别。移动台刚开机时搜索信号最强的导频信道，随后对前向同步信道(F-SYNCH)进行解码。这些同步信号用导频进行同步，并且用同样的 PN 序列加扰从而能够做到一致解调。同步信道识别基站，获取相应的系统信息，例如系统时间、广播信道和寻呼信道信息。接着，如果 cdma2000 系统向下与 IS-95 兼容，其余的细节系统信息可以通过监听前向寻呼信道(F-PCH)获取；如果不兼容，可以通过广播信道(F-BCCH)获取。

获取系统参数后，移动台能够使用某种接入方式进行注册。在基本结束方式中，首先发送一个由全零帧组成的前导，随后发送注册消息，每次发射都增加一点功率直到接收到前向公用分配信道(F-CACH)的确认消息。在重复非确认接入尝试时，使用了随机回退过程。最后，基站使用前向通用信道(F-CCCH)通报移动台的连续注册信息后，移动台就进入空闲模式，等待寻呼消息，或者使用前面讲到的接入模式请求建立一个专门用于消息传输的信道。

最后简单介绍一下 1 × EV – DV 标准。1 × EV – DV 是专门用来处理高速率分组数据的(cdma，2001[①]；Bender 等，2000[②])。我们着重讨论前向传输，因为高速率数据通常都是走这个方向，也就是因特网 Web 地址响应移动用户的方向。前向信道包括前向业务信道、控制信道，另外还有前向导频信道和前向媒体接入控制(MAC)信道。所有的信道都是时间复用的。用户分组通过前向业务信道传输，其速率是 38.4 Kbps 到 2457.6 Kbps(2.4576 Mbps)(作为对比，反向可利用的数据率从 9.6 Kbps 到 153.6 Kbps)。控制信道传输控制消息，同时也可以传输用户业务。所有零映射到一个 + 1 s 的序列上，然后用 Walsh 0 码字编码，构成导频信道，移动台利用导频信道获取最初的基站信令，这与 IS-95 以及 cdma2000 1X 等其他的系统一样。此外，移动台还用导频信号进行相位和时间的恢复，以及估计移动端的 SIR。在这个 SIR 的估计值反馈到基站后，基站用来控制传输到移动端的数据率。MAC 信道由一组 64 Walsh 信道组成。3 个子信道都通过这个信道传输，包括反向功率控制信道、DRCLock 信道(该信道在 Walsh 信道上从 5 ~ 63 开始时间复用，分配给特殊的移动终端)和反向活动信道(分配给 Walsh 信道数 4)。

前向信道中用于时分复用的信道重复传输的时隙长度为 5/3 ms。每个时隙由 2048 个速率为 1.2288 Mcps 的伪随机噪声码片流组成。图 10.6 描述了一个活动信道的时隙配置，其中业务或者控制信道占据了 4 个专门的数据位置，总共是 1600 个码片。注意图中各种信道是怎样关于时隙中心对称时间复用的，时隙被认为是对称的二等分。导频信道占据 2048 码片中的 192 个，均在两个对称的半时隙的中心。导频信道的左右都是 MAC 信道部分，见图 10.6。当需要下行传输业务或者控制信息到移动台时，总需要发送一个空余时隙。空余时隙的数据位置是空的，仍然不会填充 MAC 和导频信道。

在前向业务信道上，直接传输到一个特定移动终端的数据经常被打包到固定长度的分组内，并且以可变速率传输，这个速率是由移动台在先前反向信道的传输中针对给定分组信号选择的。这个速率取决于终端测量得到的导频信道信号的 SIR。数据传输率选择的例子将会在

---

① cdma 2001. cdma2000 high rate packet data air interface specification，3GPP2 C.S0024，Version 3.0，3rd Generation Partnership Project 2，3GPP2，December，5，http://www.3gpp2.org/public_html/specs/

② Bender，P. *et al*. 2000. "CDMA/HDR：a bandwidth-efficient high-speed service for nomadic users，" *IEEE Communications Magazine*，38，7(July)，70 – 77.

后面讲到,这个例子是基于 Bender 等(2000)的模拟结果的。各种可以利用的数据、规定的分组长度、每种数据率的调制类型等都列在表 10.4 中(cdma,2001)(控制信道只在两个最低数据率中操作)。某些数据率,尤其是值比较低时,每个分组都需要多重时隙传输。传输分组的时隙数也在表 10.4 中。当使用多时隙时,多时隙被分成 4 个时隙。这个系统指定使用 turbo 编码,turbo 码的码率也可以参见表 10.4。表 10.4 的分组长度是那些作为编码器输入的长度,这与其他 cdma2000 系统是一样的[一点不同的地方是,可变长度前导先于分组数据传输,并且占据了数据码片的第 1 部分(cdma,2001)]。

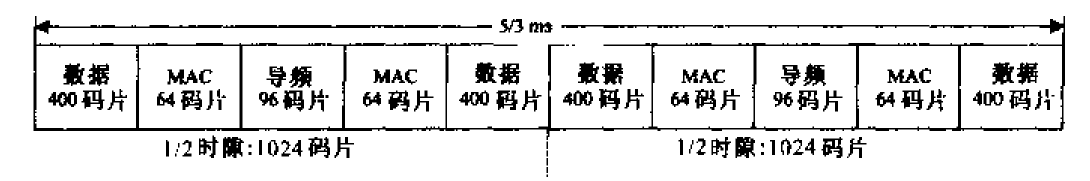


图 10.6 时隙配置,cdma2000 高速率分组数据的空中接口(cdma,2001:Fig.9.3.1.3.1-2)

表 10.4 前向业务信道参数,cdma2000 1×EV-DV:高速率分组数据系统
(引自 cdma,2001;Table 9.3.1.3.1.1-1)

数据速率 (Kbps)	分组长度(比特)	时隙/分组	码率	调制类型
38.4	1024	16	1/5	QPSK
76.8	1024	8	1/5	QPSK
153.6	1024	4	1/5	QPSK
307.2	1024	2	1/5	QPSK
614.4	1024	1	1/3	QPSK
307.2	2048	4	1/3	QPSK
614.4	2048	2	1/3	QPSK
1228.8	2048	1	1/3	QPSK
921.6	3072	2	1/3	8-PSK
1843.2	3072	1	1/3	8-PSK
1228.8	4096	2	1/3	16-QAM
2457.6	4096	1	1/3	16-QAM

需要获取各种数据率的系统配置该怎样呢?尤其是当最高数据率超过伪噪声码片速率时。多等级调制和并行传输被用来获取高数据率,见表 10.4。前向信道结构的流程图如图 10.7 所示[引自 cdma(2001)]。概括一下这个系统配置:数据分组送入编码,其输出经过混合、插入后送到调制器中进行 QPSK、8-PSK 或者 16-QAM 调制(第 5 章讨论了这些调制方法,8-PSK 将在下面讨论 EDGE 方案时简要介绍)。例如,两个连续比特合并成 QPSK 调制输入,3 个连续比特合并成 8-PSK 输入,4 比特联合成 16-QAM 输入。调制输出符号也按照需要重复,使所有数据的有效调制输出符号升到同样的值(这种情况适合较低的数据传输率)。调制输出序列符号随后分成 16 对相位(I)和积分(Q)并行流(cdma,2001)。每个流随后用速率为 76.8 K 符号/s 的 Walsh 编码。16 个分离的 I 和 Q 合并形成 1.2288 M 符号/s 的 I 和 Q 流。这两个业务流和数据前导符号流一起被导频和 MAC 信道时间复用。复合的 I 和 Q 时间复用序列随后用 1.2288 M 符号/s 的伪噪声序列编码分别编码,然后用高频余弦和正弦载波分别相乘。下面,来看一些例子:

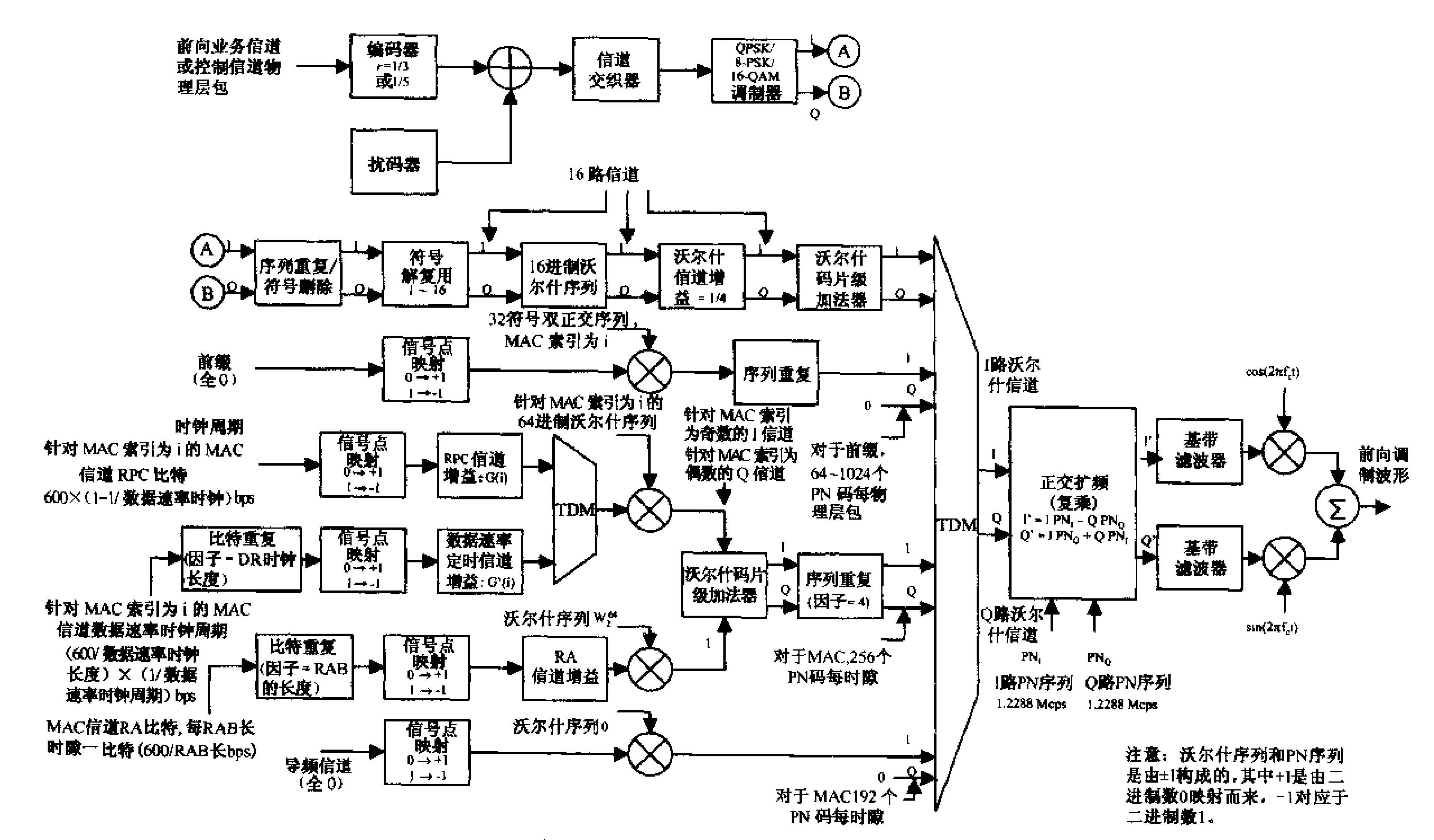


图10.7　前向信道结构（引自 cdma，2001：Fig.9.3.1.3.1-1）

例 1

最高数据率是 2.4576 Mbps,这个数据率对应的分组大小是 4096 比特,以 5/3 ms 时隙传输,如表 10.4 所示。输出使用 1/3 码率的编码器,速率是 3 倍于这个速率,达到 7.3728 Mbps。这个比特流通过 16-QAM 调制,一次收集 4 个比特用来调制,也就是说调制输出符号率是 1.8432 M 符号/s,或者 3072 符号每 5/3 ms 时隙。这个符号流现在分离到 16 个 Walsh 编码流的并行对中。

例 2

编码器输入数据率是 1.8432 Mbps,相应地,在编码输入端是 3072 比特数据分组每 5/3 ms 时隙。因为使用 8-PSK 调制,又有了一个调制输出率 1.8432 M 符号/s,或者以 3072 符号每时隙分离到 16 个 Walsh 编码流的并行对中。然而,同样的传输输出功率,2.4576 Mbps 数据流由于使用了 16-QAM,在接收端需要更高的 SIR,而 1.8432 Mbps 的数据流使用的是 8-PSK。

例 3

最后,考虑数据率是 1.2288 Mbps 的情况,从表 10.4 中可知,有两种选择。一种情况是一个 2048 比特的分组在一个时隙内传输,用 QPSK 调制。另一种情况是,用两个时隙传输一个长度是 4096 比特的分组,使用 16-QAM 调制。由于时隙间隔,在这种情况下会引起额外的延时。接着谈表 10.4 中其他的例子,最低数据率是 38.4 Kbps,1024 个比特数据分组需要 16 个时隙来传输,每个分组之间有 4 个时隙间隔。该例子中,延时例如传输和接收一个包的时间,但并不包括处理和分组缓冲的时间,并不是很大。但是由于位宽很大,并且使用 QPSK 而不是 8-PSK 或 16-QAM,所以接收端所需的 SIR 比高比特率传输数据要小。

使用这个高速率的分组数据系统,用户当然希望能带来更高的速率。究竟能达到什么样的数据率,怎样把数据率和延时带来的问题做折中? Bender 等(2000)做了一个 cdma2000 数据网络仿真,其特征大部分都和现在的 cdma(2001)相同(编码率与表 10.4 所示的略有不同,例如,假设低比特率系统使用 1/4 码率而不是 1/5 码率的编码,更高比特率使用的是 1/2 而不是 1/3 码率的编码。除了最高速率的 1843.2 Kbps 和 2457.6 Kbps 外,别的数据率都用 QPSK 调制,而 1843.2 Kbps 使用 8-PSK,2457.6 Kbps 使用 16-QAM 调制。注意表 10.4 中的 921.6 Kbps 也使用 8-PSK)。在仿真实验中,研究终端的接收 SIR,包括干扰参数的估计,蜂窝内、蜂窝间以及有热噪声时的估计。这个 SIR 估计值由终端转化成最大数据率,用来估计给定的分组数据错误率。这个最大数据率随后会传到基站,就如先前讨论过的一样。

Bender 等(2000)在表 2 中列出了该项研究中得到的为了维持 1%分组错误率所需的 SIR 值。有趣的是,在数据率是 38.4 Kbps 到 1228.8 Kbps 这个范围内时,SIR 几乎是线性增加的。38.4 Kbps 数据率需要的 SIR 是 - 12.5 dB。在 38.4 Kbps 到 1228.8 Kbps 这个范围内,数据率每增大一倍,SIR 就需要增大约 3 dB(注意当数据速率比值为 32 dB 或 15 dB 时,全部 dB 差为 15.5 dB)。当速率增至 2.4576 Mbps 后,SIR 需要增大到 9.5 dB。Bender 等同时画出了 SIR 的累积分布函数图(cdf),该图表明 SIR 有 50%的时间小于 0 dB,有 80%的时间小于 5 dB(也就是说,SIR 大于 0 dB 的概率是 0.5,大于 5 dB 的概率是 0.2)。结合 SIR 的 cdf 测量值,以及给定数据率时所需要的 SIR 值,每个数据率能够被传送的概率就可以计算出来了。该结果也被 Bender 等制成表格,并且把数据率画成柱状图。概率最大的数据传输率是 307.2 Kbps,概率是 0.24。其他数据率大体上是以 307.2 Kbps 为中心对称分布的。例如,153.6 Kbps 的概率是 0.08,而 921.6 Kbps 的概率是 0.09。所有传输速率以各自出现的概率加权,再求和所得的值就是平均系统吞吐量,这个值在 570 Kbps 左右。

有可能超过这个平均吞吐量的值吗? Bender 等建议采用基于限制延时的技术。注意,这

个简单的计算平均系统吞吐量的方法完全忽略了延时。而延时会随着传输率变化而变化。在表 10.4 中,传输给定大小分组的时隙数是和数据率成反比的。由于低的 SIR 会带来更长的传输延时,所以用户不得不使用低的传输速率。Bender 等建议采取一定的策略,把最大和最小延时的比固定在某个具体的值上。他们同时介绍了在给定延时率的情况下,如何使平均吞吐量最大化。增加延时率会增加平均吞吐量。如果用户需要有同样的延时率,不管数据率如何,系统仿真的平均吞吐量能达到 360 Kbps。如果延时率增加到 8,则平均吞吐量能增加到 750 Kbps。

前向信道中适当传输策略的讨论只是在本章开头部分讲到的如何分配系统资源的一个例子,资源分配问题又称为调度问题。我们将在第 11 章着重介绍这个问题,现在以多重接入的一般问题为例:系统资源是怎样合适地分配前向和后向的移动台用户和不同类型的业务,来获得比较好的目标的? Andrews(2001)[①]发现了一种比较好的方法,并且将其直接应用到 cdma2000 高速率传输系统中去了。该方法讨论了吞吐量优化调度的原理,使用别的性能目标函数来替代上面讲到的延时率。

10.3　2.5/3G TDMA:GPRS 和 EDGE

前面讨论了 3G CDMA 系统,现在来介绍使用 TDMA 技术的 3G 系统,重点介绍起源于 GSM 并且和 GSM 兼容的系统。这些基于 TDMA 的 3G 系统通常也考虑 2.5G 系统:尽可能地减少 GSM 结构的变化,而能给 GSM 提供分组交换能力,并且数据传输率尽可能地快。有两种标准 GPRS(通用分组无线业务)和 EDGE(全球演进式增强型数据速率)可以满足这个要求[②]。GPRS 有两个基本部分组成:核心网络部分和空中接口标准。EDGE 用来增加空中接口的比特率,IS-136 系统将吸收这些新技术和结构(IEEE,1999)[③]。GPRS 和 EDGE 的集合有时又称为 EGPRS,即增强的 GPRS。首先介绍 GPRS,随后再介绍 EDGE。GPRS 核心网络部分使用层结构,基于因特网的 TCP/IP 协议在其中起着关键性作用。因此,让我们花点时间,简要回顾一下分组交换系统中的层结构,并且主要以 TCP/IP 为例(WCDMA 和 cdma2000 也在分组交换结构中使用了层结构,但是由于之前一直在介绍空中接口部分,所以一直都没有讨论过层结构这个概念。在第二代蜂窝系统中,层用来描述各种消息传输,但是在第 8 章中并没有讨论过这个问题,以免引起不必要的混淆)。

GPRS

正如前面所讨论的 3G CDMA 系统,GPRS 技术和结构在设计时就考虑了基于分组的应用。在本节中,简要介绍它们在 3G 基于分组系统中的应用和需求。例如,一系列数据传输类型都是这些应用的结果:突发数据业务,暗示投递数据的时间需求比连续突发的时间短;频繁传输小容量数据;使用 Web 地址时,偶尔会传输大容量数据。这些应用使 GPRS 具有特殊的属性

---

① Andrews, M. *et al*. 2001. "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, 39, 2 (February), 150 – 153.

② 对"GSM"和"GPRS"中的首写字母缩略词 G 的另一种解释可在文献中见到。在第 8 章 GSM 代表全球移动通信系统(Global System for Mobile Communication),最初是由法语词组 Groupe Speciale Mobile 演变而来。字母缩略词可能一直不变,但它们的意思却可能改变!

③ IEEE. 1999. "The evolution of TDMA to 3G," *IEEE Personal Communications*, 6, 3(June).

(GPRS,1998[①];Kalden 等,2000[②])(前面讨论过的其他 3G 系统也借鉴了这种属性)。它们包括有效使用射频资源,可以比处理电路交换语音业务时间更短;分组传输的有效性;与其他分组网络(PDN,如因特网)无缝连接的能力;最后,对所期望的服务质量(QoS)的规定。

我们在讨论 WCDMA/UMTS 时,曾定义了一系列业务类型,每种类型都是以诸如延时、信息实体之间的时变、数据完整性(错误率)等 QoS 作为特征的。GPRS 中的 QoS 定义成一系列的服务特征(GPRS,1998)[③]。用户订阅具体的在一个契约周期内的用户描述,这些用户描述包括一个 QoS 描述。这个描述里面的 QoS 属性包括服务优先级、可靠性、延时和吞吐量(GPRS,1998)。服务优先级分为 3 个等级:高优先级、中优先级和低优先级。低优先级应用于 WCDMA 背景的类型或者因特网中那些尽力而为的业务,这些业务对信息传输时间的要求并不是很严格。投递一个用户信息单元的可靠性有 4 个参数:数据单元的丢失率、重复率、乱序的概率(数据单元不是按照发送顺序到达的)和数据被破坏的概率。第 1 类需要这些概率的数量级为 10^{-9}。这些类型的业务对错误率是敏感的,没有错误校正能力,对错误的容忍能力是有限的。可靠性第 2 类规定的这些参数的值为中等程度,大概是从 10^{-4} 到 10^{-6}。这些类型的业务对错误也很敏感,但是有一定的纠错能力,对错误的容忍能力也比较高(GPRS,1998)。最后,第 3 类应用于对错误并不敏感的业务,这种类型的业务的丢失率可以达到 10^{-2},数据单元被破坏的概率也能维持在这个值,数据重复率和乱序的概率都可以达到 10^{-5}(GPRS,1998)。

GPRS 有 4 种延时类型,请参见表 10.5,该表引自 GPRS 的表 5(1998)。"延时"指的是传输数据通过 GPRS 网络所需的时间,并不包括通过外部网络(例如,因特网)到 GPRS 网络的时间。这个时间包括接入无限信道,接着上传一个数据单元所需的时间,穿过 GPRS 网络的时间,调度时间,以及下行传输一个数据单元到移动台的时间。表 10.5 的延时时间包括平均传输延时和 95% 延时,并且有两种不同的数据单元,尺寸大小分别为 128 字节和 1024 字节。注意这些延时值的大小,显然不能承担实时交互的语音业务,语音业务端到端的最大延时大概是 0.1 s。但是,对于某些类型的数据业务,平均延时到分钟级也是可以接受的。

表 10.5 GPRS 延时类型(引自 GPRS,1998;Table 5)

	最大延时(S)			
	数据单元大小为128个八位字节		数据单元大小为1024个八位字节	
延时类别	平均延时	95 个百分点	平均延时	95 个百分点
1	<0.5	<1.5	<2	<7
2	<5	<25	<15	<75
3	<50	<250	<75	<375
4			(不确定)	

① GPRS. 1998. "Digital cellular communications system(Phase 2 +); general packet radio service description," Stage 1, EN 301 113 v6.1.1 (1998 - 11) (GSM 02.60 version 6.1.1 Release 1997), European Telecommunications Institute (ETSI), Sophia Antipoli, Valbonne, France.

② Kalden, R. et al. 2000. "Wireless internet access based on GPRS," *IEEE Personal Communications*, 7, 2(April), 8 - 18.

③ GPRS. 1998. "Digital cellular communications system(Phase 2 +); general packet radio service description," Stage 1, EN 301 113 v6.1.1(1998 - 11) (GSM 0260 version 6.1.1 Release 1997), European Telecommunications Institute(ETSI), Sophia Antipoli, Valbonne, France.

在用户描述中最后一个 QoS 参数是用户数据吞吐量。这个参数代表用户请求的数据吞吐量。它包括两个协商参数:最大比特率和平均比特率(GPRS,1998)。突发业务传输的平均速率包括没有数据传输时的时间。例如,用户传输数据经过协商后的最大速率是 384 Kbps,但是只用了 5%的时间,它的平均速率是 19.2 Kbps。分配给用户的最大比特率显然受限于系统所能够达到的最大比特率。因为 GPRS 允许 GSM 分组也能分组交换数据,所以它也采用了 GSM 的 8 时隙重复帧结构。电路交换的 GSM 每帧给用户分配一个时隙,分组交换 GPRS 则与之不同,用户最多可以分配到 8 个时隙以增加数据吞吐能力(这样降低了接入系统的用户数,但是这是一帧中的时隙,而电路交换的时隙还要更长一点)。每帧最大只能分配 8 个时隙,这就限制了吞吐量所能达到的极限。

在现实应用和市场需求中,有的移动用户需要同时使用 GPRS 分组交换和 GSM 电路交换业务;而有的只需要使用 GPRS 或者 GSM 业务。由于这个原因,定义了 3 种不同的 GPRS 移动台类型(GPRS,1998)。类型 A 同时支持 GPRS 和 GSM 业务,"同时"是指在同一个时隙内两种业务都能够被使用。因此,这种类型每帧的时隙的最小值必须是明确可利用的[①]。类型 B 不支持同时传输两种业务,但是它能够接收这两种业务的呼叫,使用的业务能被自动地选择。最后,类型 C 只支持其中某一种业务,但是移动台能够配置成另一种类型。

接下来着重讨论 GPRS 所能支持的与 QoS 相关的应用类型,我们也将讨论移动台所支持的不同类型。现在介绍处理这些应用所需要的 GPRS 结构,提供所需要的服务质量描述。在 GSM 核心网络结构中,GPRS 需求的执行会有所不同,GSM 空中接口的消息结构传输也会有所不同。首先讨论 GPRS 核心网络,接着讨论空中接口中的消息传输。为了建立 GPRS 网络,在 GSM 系统结构中加入了两种新的基于分组的支持节点。它们是服务 GPRS 支持节点(SGSN)和网关 GPRS 支持节点(GGSN)。GPRS 移动台利用 SGSN 连接基站,向其发送数据包或者接收数据包。因此,这与电路交换 GSM 移动交换中心(MSC)类似。SGSN 之间用公共地面移动网络(PLMN)连接,或者通过 GGSN 连接到外部分组数据网络(如因特网)。SGSN 使用分组路由,并且与电路交换的 MSC 一样,负责移动管理、位置管理、用户认证,以及其他在服务区内的功能需求。注册在 SGSN 服务区域内的所有 GPRS 用户的位置信息都存储在 SGSN 中,位置信息包括本地位置寄存器(HLR)和访问位置寄存器(VLR)。

图 10.8 描述了 SGSN 和 GGSN 在各种数据网络结构中是如何应用的。两个 SGSN 各自控制着一定数目的基站,并且用公共地面网络(PLMN1)互联着。作为特例,SGSN 能够直接互联(为了描述简单,一个 SGSN 只控制一个基站)。SGSN 通过 GGSN 连接外部数据网络(PDN),如图所示,PLMN2 用来连接公共数据网络。为了使图简单,只用了 3 个数据移动台 M1 ~ M3。注意,其他数据移动台也许会借助基于地面的数据网络连接到其他数据移动台、数据用户、各种分组交换数据网络中的数据库。其他可能的配置允许在移动台或者固定数据终端和数据库之间使用多重路由路径(Bettstetter 等,1999)[②]。

---

① 这种类型的配置暴露了在已存在并广泛布网的 GSM 系统与这里讨论的 GPRS 系统互相作用时的问题。包括对于不同的传播特性,3G GPRS 系统使用了更高的频带,这导致在建筑内部更弱的信号穿透力。这也将导致当用户穿过建筑物时,在 GPRS(或者 WCDMA)与 GSM 系统间移交的增加。另一个显著的问题是要保持移动终端的低成本,并要同时支持 GSM 和 GPRS 系统。在需要解决这些问题的基础上,产生了类型 A 双传输模式(DTM)。这些问题的提出以及使用类型 A 双传输模式(DTM)进行解决的细节,详见 Pecen 和 Howell(2001)的著作。

② Bettstetter, C. *et al*. 1999. "GSM Phase 2+ general packet radio service GPRS: architecture, protocols, and air interface," *IEEE Communication Surveys*, 3rd qtr.

层结构回顾

数据单元是如何在分组交换网络内传输的，例图 10.8 所示，从一个移动用户到一个目的端(如 Web 地址或者另外一个移动用户)？分组交换网络使用层结构来保证数据单元以某种恰当的形式到达所期望的目的地(Kurose 和 Ross，2002[①]；Schwartz，1987[②])。而每层的操作都是通过已经确定的协议进行的("协议"这个词明显是来自外交界的，代表人们之间互相通信所需的程序)。某一层的协议用来在逻辑上连接同样层的另一端。接着我们将看到，GPRS 的核心网络使用层结构来保证网络内和网络间的正确通信。但是在介绍 GPRS 结构和协议之前，先回顾一下层结构的概念。以因特网层结构为例。

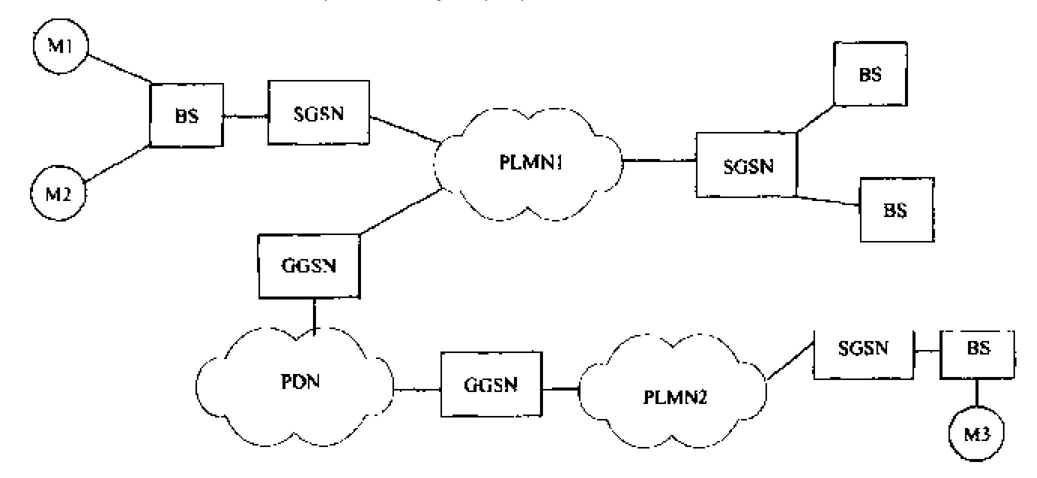


图 10.8 简单的 GPRS 网络配置

一般说来，下面的层是为上面的层提供服务的。图 10.9 是一个使用因特网结构的简单例子，图 10.9(a)显示了一个网络的一部分，有两个用户主机、终端或者计算机，通过两个路由器连接另一台主机。每个路由器另外的输出连接了其他路由器或者主机。图 10.9(b)显示了图 10.9(a)所示的设备所对应的各层结构。每个主机有 5 层结构，路由器只有 3 层。每层相应的协议一般都是用软件执行的，只在最底层用硬件实现。虚线表示设备间各层的逻辑连接。最底层的实现表示进行物理连接，例如用无线电、电缆或者光纤连接，通常都是按比特传输的。

用户在应用层输入将要被传送的信息。这些信息以信息数据单元的形式组成，通常称为信息的有效载荷，必须通过网络路由到达目的端。应用层下面的层保证数据单元到达所期望的目的地。首先考虑应用层下面的一层传输层，在 Internet 协议组里又称之为传输控制协议(TCP)。该协议主要用在非实时应用中。语音等实时应用则是使用用户数据包协议(UDP)。引用 TCP 只是为了简单起见，该协议为应用层以上提供数据序列传输、错误纠正、向网络顺序投递等服务。错误纠正是利用自动重复请求(ARQ)实现的，第 7 章曾简要地提到过(在图 10.9 的发送系统中，当开始传输一个序列时就设置定时器。如果接收端主机正确接收到这个消息，它就返回一个确认给发送端。如果接收端收到的消息不正确，它就扔掉这个消息。如果发送端在定时器定时范围内收到确认信息，它就假设接收端已经正确接收这条消息了，因此就从存储空间删掉这条消息。如果在定时器过期的时候还没有收到确认信息，发送端就重新发送这条消息)。

① Kurose，J. F. and K. W. Ross. 2002. *Computer Networking*, Boston, MA, Addison-Wesley.

② Schwartz, M. 1987. *Telecommunication Networks: Protocols, Modeling, and Analysis*, Reading, MA, Addison-Wesley.

TCP还提供流控制操作，来保证数据顺序传输且接收到的消息并没有溢出接收器的缓存。

在图10.9(b)中，网络两端主机的应用层和TCP层用虚线连接表示一端的层在逻辑上是与另一端对应的层连接的。这个层上的端到端的操作，不管是发送数据还是接收数据，都不需要知道具体的网络结构是怎样的。TCP提供给应用层服务的目标是把消息放入发送端管道，保证它们正确传输，并且在接收端顺序到达。但是，消息实际上并不是通过这些“管道”传输的。携带这些消息的数据单元必须通过网络才能路由到另一端。这个路由功能是由TCP下面的网络层提供的，在因特网中，网络层依赖于因特网协议(IP)执行。IP安装在图10.9(b)的每台主机和路由器中，系统之间用虚线连接。注意，虚线表示这是逻辑上的连接，并不是真正的物理连接。

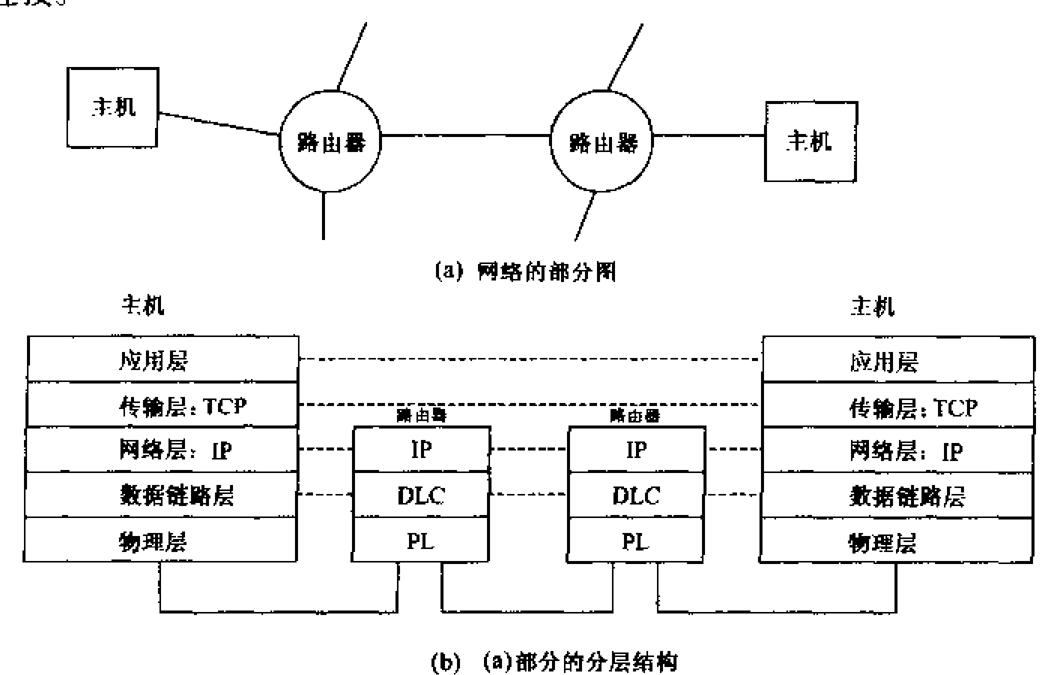


图10.9 因特网TCP/IP层结构

数据之间真正的物理传输必须沿着相邻设备之间的传递才行，如图10.9(a)所示的实线部分。数据链路层的数据链路协议保证携带数据的比特块能够正确地到达链路的接收端(这一层分为两个子层，将在下面的GPRS层结构中讲到)。数据链路协议的一个作用就是识别给定数据单元的开始和末尾。协议经常设计成在链路的另一端数据单元的接收是否有错误、消息的接收是否有错误，或者有时候简单地告知它的上层检测到错误。数据链路层有个子层用来接入过程，例如，Aloha协议就被使用在无线通信的无线电链路连接中。数据链路层是逻辑连接的，在图10.9(b)中用虚线连接。

物理层在最底层，是设备与系统的实际物理连接，如图10.9(b)的实线部分所示。这个层的传输比特组成了数据单元。无线电传输就是物理层传输过程的一个例子。各种传输技术例如PSK，QAM和CDMA都是物理传输过程的一部分。物理层的目标是保证比特字节能够沿着链路传输，这个链路可以是无线链路、双绞线或者光纤。

注意一直使用“数据单元”来表示各层中的二进制块，本书的其他部分也用这个词来代表二进制块。在实际过程中，应用层发送一条消息，各层中携带这条消息的二进制块都是不同的。消息可能太长了，以至于用一个数据单元无法传输；这样就需要把消息分割到各个单元

中。在第8章,我们已经见过这样的例子了。相反,如果消息太短,则传输就不很经济。许多条消息可以联合成一个更大的数据单元。最重要的是,消息在逐层往下传递时,每层的协议都会在消息前面加上自己所对应的"头"。这些"头"携带着各层的信息,接收端所对应的层需要这些信息。在某一层,"头"加上从上一层接收来的数据单元称为协议数据单元(PDU)。在接收系统,协议数据单元移除头信息,并向上层传递。随着协议数据单元在层间上下移动,以增加或删除头信息,会使其发生变化。由于这个原因,每层的协议数据单元都有自己特定的名字:应用层协议数据单元(A-PDU)、传输层协议数据单元(T-PDU)、网络层协议数据单元(N-PDU)以及数据链路层协议数据单元(DL-PDU)。我们可以把这些PDU看成是在两台主机对等层之间用相应的协议进行逻辑连接之后进行的流动。而实际的物理传输只发生在物理层,原始数据在加上各种头信息后下移到物理层,由物理层进行实际的传输。

在图10.10显示了各个协议数据单元逐层往下移时,在各层加上头的过程。DL-PDU包含了所有各层的"头"并在物理层进行传输。而发送主机则认为它是在以应用层协议数据单元的形式传输数据,接收主机也只看到A-PDU。下面的层在消息上加上各自的头信息,来保证原始消息能够在网络内进行正确的路由和投递,如图10.10所示。下面我们在讲述GPRS层结构的时候会涉及到具体的细节。

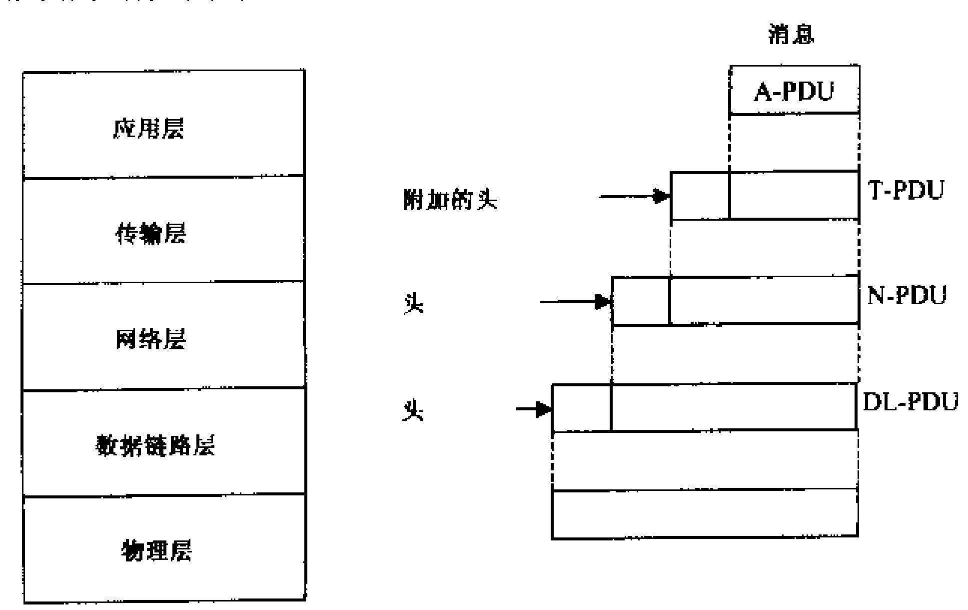


图10.10 层和协议数据单元

迄今为止,我们一直在讨论PDU,那"分组"可用在哪里呢?严格来说,分组和网络层协议数据单元(N-PDU)是同义的。因此,在因特网结构中,IP层处理分组并且处理分组在网络中的路由,又称为IP分组。而在非正式的使用中,其他层的数据单元也常称为分组(例如,尽管以太网包括数据链路层的下面一层媒体接入层(MAC层)和物理层,以太网数据单元有时候也称为以太网分组)。只要讨论时不会引起混淆,这些说法是允许的。但是当同时讨论不同层的数据单元时,就得注意不要混淆了。

GPRS层结构

回顾了层结构之后,我们讨论GPRS网络中的层结构。其中将讨论更多协议的细节,这些协议用来定义移动台和基站之间空中接口的交互。

图 10.11 简单描述了 GPRS 的层结构,主要是一个移动台及它所对应的基站、BS 所关联的 SGSN 以及把 SGSN 连到外部分组数据网络的 GGSN(Kalden 等,2000;Cai 和 Goodman,1997[①])。虽然实际的比特传输发生在各种物理层,但是为了叙述简单,每一层之间的连接都用实线表示。除了物理层,其他的连接都是逻辑连接。我们着重讲带有因特网或者其他配置 IP 的分组数据网络的基于 TCP/IP 的通信。GPRS 也提供了其他基于分组网络的业务,如 X.25(Kalden 等.,2000;Cai and Goodman,1997;Bettststter 等.,1999)(X.25 协议在 Schwartz 于 1987 年写的著作中进行了讨论)。首先注意,由于目标主机并没有在图 10.11 中画出,这幅图中的应用和 TCP 层并不是终端。GGSN 中的 IP 层支持这些层的终端,因此 GPRS 网络在分组网络中被认为只需要一跳就能到达。所以,逻辑上 IP 分组通过 GPRS 核心网络从移动台传输到它所连接的 GGSN。子网相关的汇聚协议(SNDCP)允许在一个逻辑连接上复用 IP 分组,如果需要的话,我们就把这些分组段分成更短的数据单元,并且提供压缩能力。只有当 SNDCP 需要变化时,另一个网络协议才被用来替代 IP。图 10.11 中其他的协议不会受到影响。注意 SNDCP 层在移动台和 SGSN 之间通过基站提供了一个逻辑连接。虽然 SNDCP 层结束于 SGSN,但是 SGSN 作为一个中继站可用来传递 IP 分组到 GGSN,GGSN 可把 SGSN 连接到 GPRS 核心网以外的分组数据网络中。

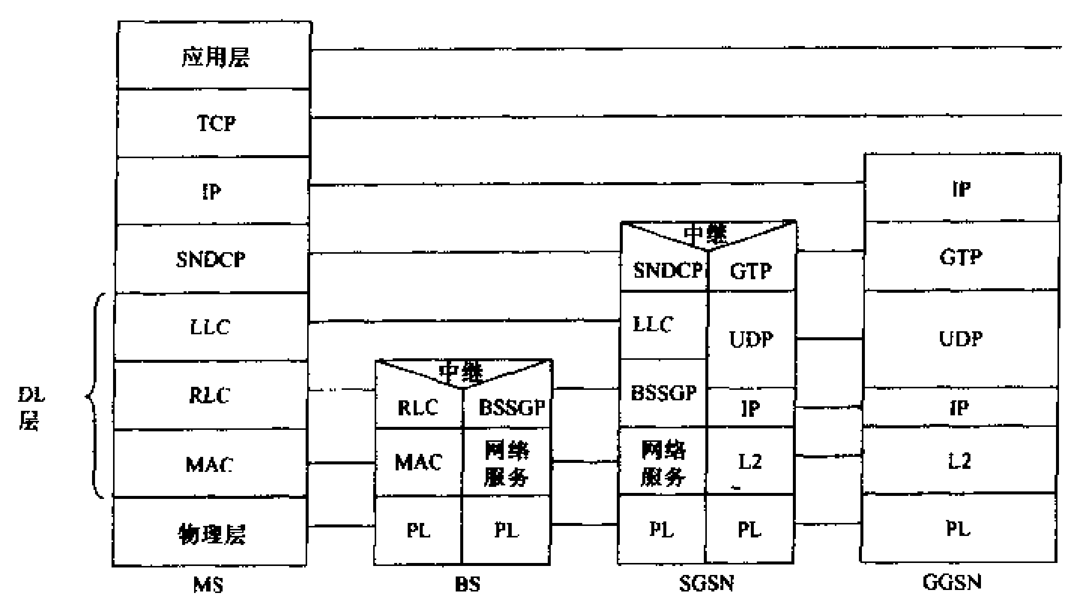


图 10.11 GPRS 层结构例子

在 SNDCP 层下面,且给 SNDCP 提供服务的层是数据链路层。在图 10.11 中,这个层有两个子层组成:逻辑链路控制子层(LLC)和无线电连接控制(RLC)与媒体接入控制(MAC)子层。把数据链路层分割成 LLC 和 MAC 子层在分组交换网络中是很常见的,IEEE 800 系列的第 2 层数据链路层协议曾作为典范广泛引用(Kurose 和 Rose,2002;Schwartz,1987)(局域网 LAN 也就是以太网都是用层 2 协议。以太网和 IEEE 802.3 是同义的。在第 12 章中讨论了两种 WLAN 标准 802.11 和 802.15)。RLC 和 MAC 子层的联合只出现在 GPRS 中。注意,LLC 子层和 SNDCP 层一样,在移动台和 SGSN 之间提供连接。出现在基站的 LLC-PDU 直接传递到 SGSN。RLC/MAC 子层逻辑连接在移动台和基站。LLC 子层给上面的 SNDCP 提供了一系列服务,包括流和顺序控制、错误控制。当需要时,RLC/MAC 子层的 RLC 协议可提供给 LLC PDU 的段,也可提供

① Cai, J. and D. J. Goodman. 1997. "General packet radio service in GSM," *IEEE Communications Mag*., 35, 10(October), 122–131.

RLC/MAC PDU 的错误控制。错误控制能够使用 ARQ 技术纠正错误,或者仅仅发现错误并告知上层(这些技术在第 7 章中提到过)。

媒体接入控制(MAC)协议给移动台接入基站提供接入控制,它使用本书前面已经讨论过多次的时隙 Aloha 接入技术。在需要时,它还负责给移动台提供多个时隙。MS-BS 无线电接口的物理层实际上由两个子层组成。上层是物理链接子层,使用卷积码来进行前向纠错,它同时还负责空中接口的信号质量检测和功率控制。下面的子层是物理无线电频率层,提供实际频率分配和通过空中接口传输信号的调制特征,这个子层与 GSM 中的一样。在讨论完图 10.11 中 GPRS 各层和协议之后,我们会更详细地讨论空中接口更底层的协议。

(在本章前面所讨论的蜂窝系统以及第 8 章所讨论的电路交换系统都着重讨论了空中接口的物理层,并在讨论接入过程和逻辑消息传输的时候暗含着讨论了 MAC 层。我们当时没有评价这些系统中的协议层,以免引起混淆。现在读者可以重新读一下前面的章节,思考一下设计这些系统时层的概念是怎样被使用的)。

在继续讨论图 10.11 其他协议层之前,最好知道在各种层中 PDU 是怎样形成的。最底层的 PDU 通过空中接口在 MS 和 BS 间传输。各层的 PDU 请看图 10.12(Kalden,2000)。有些层的 PDU 的段纯粹是任意选择的,用来给这个过程提供例子。移动端产生的应用 PDU 分割成两部分,以形成 TCP 有效负载(应用 PDU 没有规定最大长度,TCP-PDU 的最大长度是 64 000 字节)。分割后的两个段各自都有标准的 TCP 头,它们通过加上符号 H 形成 TCP-PDU。这个 TCP-PDU 可形成 IP 有效载荷,IP 再加上它的头符号 H。所形成的 IP-PDU 或者网络分组就由 GGSN 在逻辑上进行传输,如图 10.11 所示。作为例子,IP 分组又被分成两段,每段加上 SNDCP 头从面形成 SNDCP-PDU(SNDCP-PDU 长度限制在 1560 字节,而 IP 分组包可以长达 64 K 字节)。SNDCP-PDU 传到 LLC 协议层,再加上相应的头。最后,LLC-PDU 被分段成多个 RLC/MAC-PDU 后,用物理层所提供的服务在空中接口中进行最后的传输:我们将在下面的 GPRS 空中接口部分看到,RLC/MAC-PDU 的长度限制在 54 字节,并且包括了各层加上去的头。在开始讨论图 10.11 时,LLC-PDU 通过基站从移动台传到 SGSN。这意味着由 LLC-PDU 分割而成的 RLC/MAC-PDU 必须在基站重新组合,去掉 RLC/MAC 头之后恢复成 LLC-PDU。LLC-PDU 随后原样传给 SGSN。RLC/MAC 的头信息为基站执行 RLC/MAC 协议提供了必要的信息。

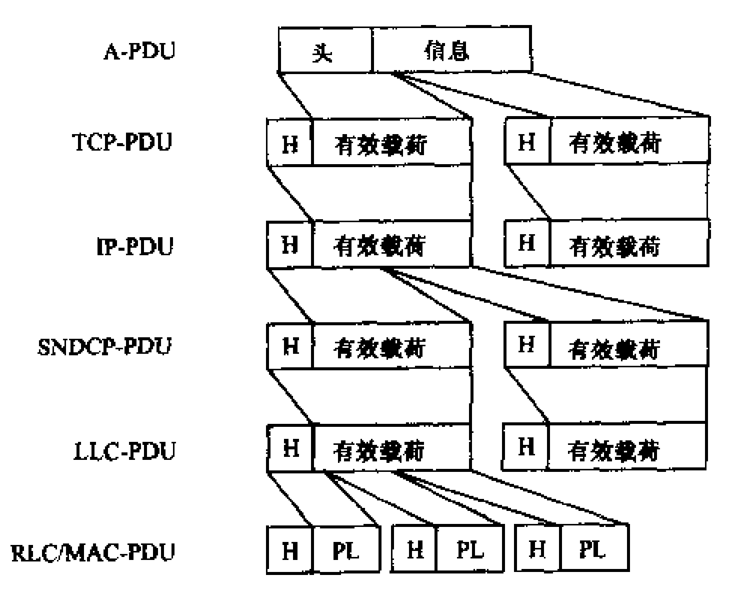


图 10.12 GPRS 空中接口 PDU 的产生(引自 Kalden 等,2000:Fig.3)

现在回到图 10.11,简单描述一下图中其他的协议层。图 10.8 显示了一个很简单的 GPRS 网络配置。在为多个基站提供服务的 SGSN 之间,通过公共地面移动网络(PLMN)进行互联,并且通过 GGSN 和外部分组网络进行连接。GPRS 结构提供了基于 IP 的 PLMN,也就是说,所有通过 PLMN 的必须使用 IP 以及与 IP 相关的协议。但是,SGSN 和 GGSN 之间的信息传输连接,逻辑上是一跳就可以到达,如图 10.11 所示。在 SGSN 和 GGSN 之间的 PLMN 路由必须被 IP 分组隐藏起来,这个 IP 分组是从移动台传递到连接着外部分组网络的 GGSN。这个通过 PLMN 的隐藏的路由又称为"隧道":IP 分组在 SGSN 经过封装,并加上携带有目的 GGSN 地址的头部,然后用 PLMN 相关的 IP 协议通过 PLMN 路由到 GGSN。GPRS 隧道协议(GTP)所使用的隧道画在图 10.11 中。逻辑上,这个协议认为 SGSN 和 GGSN 是直接互联的,而物理上可能需要跳很多次才能到达。这个直接逻辑连接是通过 UDP 实现的,因为基于 IP 分组的数据网络没有提供可靠的网络传输(图 10.11 中的 TCP 层保护移动公共数据网络的连接。X.25 有可靠的数据连接,因此使用它在移动台和 GGSN 之间提供网络服务来替代 IP。TCP 用于服务 GTP 层,逻辑上将 SGSN 和 GGSN 连接来替代 UDP)(Bettstetter 等,1999)。PLMN IP 层用于给封装的信息通过 PLMN 提供路由,如图 10.11 所示(必须搞清楚的是,移动的 TCP 和 IP 协议不能与 PLMN 内部的 UDP/IP 协议混合。移动端产生的以外部分组网络为目的地的 IP 分组使用内部的 PLMN 协议通过 PLMN 进行传递,这个协议就是 UDP 和 IP)。最后,注意到在给定的 PLMN 中没有具体指明数据链路层和物理层协议,因此在图中这两层称为 L2(层 2)和 PL。

图 10.11 中所需要讨论的最后一个协议是 BSSGP。基站子系统 GPRS 协议(BSSGP)在基站和 SGSN 之间提供了一个逻辑连接,用来传递必要的路由和 QoS 信息。两者之间的连接由一些网络服务和与这些网络服务相关的物理层协议所提供。

移动台和 GPRS 网络建立连接的过程,以及与外部分组网络(如因特网)建立起一个对话,我们把它们作为图 10.11 中层结构如何运作的例子。该过程还包含对话所需的 QoS 谈判的过程。为了连接到网络,移动台首先必须在 SGSN 中注册,这个过程称为 GPRS 连接(Bettstetter 等,1999;Koodli 和 Punskari,2001[①])。一旦连接成功,移动台便开始激活一个称为分组数据协议(PDP)报文的过程。在基于 IP 的环境中,假设这个 PDP 报文与在移动台和 GGSN 之间基于 IP 的虚连接相对应(Bettstetter 等,1999;Koodi 和 Punskari,2001)。在 GPRS 连接的过程中,移动用户认证是有效的,用户描述信息从 HLR 拷贝到 SGSN 中。从图 10.11 我们注意到,在这个过程中移动台和 SGSN 之间所交换的消息只需要结构上的更低层即可。

为了激活与外部网络进行会话而建立的 IP 类型 PDP 报文,需要先确定移动 IP 地址(如果事先没有给移动台分配的话)和协商会话所需要的 QoS(移动家庭网络的 IP 也许是永久分配给移动台的或者地址是在会话建立的时候动态分配的。后面一种情况的地址由 GGSN 分配)。激活过程需要在移动台、移动台对应的 SGSN 和 GGSN 之间传输一系列请求、响应等消息(Bettstetter 等,1999)。在连续报文激活的过程中,这些消息在图 10.13 中得以描绘(Bettstetter 等,1999:Fig.4)。移动台首先发送一个激活 PDP 报文请求消息给它的 SGSN。这个消息包括 PDP 类型(如 IPv4)、IP 地址(如果先前已经分配了)、QoS 描述请求以及其他消息参数(如果 IP 地址没有被分配给移动台,地址域空白)。这个消息使用图 10.11 所示的 SNDCP 层发送,

---

① Koodli, R. and M. Punskari. 2001. "Supporting packet data QoS in next-generation cellular networks," *IEEE Communications Magazine*, 39, 2(February), 180 - 188.

SNDCP逻辑上连接移动台和SGSN。消息的物理传输需要通过基站传递，如图10.12所示，创建和连接LLC头，然后把LLC-PDU分割成多个RLC/MAC-PDU，每个都加上MAC/RLC头，沿着MS-BS空中接口传输，如图10.12所示(RLC/MAC-PDU如何在空中接口传输将在下一节讲到)。用户认证由SGSN进行，如果用户通过认证，SGSN将发送一个PDP报文请求消息给GGSN。这个消息同时包括移动IP(PDP)地址(如果移动地址还未被分配，该栏空白)、一个QoS描述域以及其他参数。这个消息使用GTP和支持GTP下面的层，通过SGSN传输到GGSN。这个报文激活过程也许还进行了QoS参数请求的协商。QoS属性最终达成一致，并且必须要有可以利用的网络资源。如果有可以利用的网络资源，那么协商成功之后GGSN发送一个PDP报文响应SGSN，如图10.13所示。然后，SGSN再向移动台发送一个激活PDP报文的接受消息。如果还没有给移动台分配IP，GGSN就分配一个IP给移动台，并用PDP报文响应告知SGSN。在SGSN发送给移动台的激活PDP报文中，其地址域就填充了这个地址。SGSN把协商成功的QoS映射成RLC/MAC无线电优先级，并把这个优先级通过传输给移动台的接收消息告知移动台，作为随后上行连接接入请求的优先级。由基站返回的请求决定移动接入的优先级。如果GPRS核心网络具有相应的功能的话，QoS描述协商还将被SGSN和GGSN映射到IP的QoS参数中(Koodli和Punskari,2001)。移动台发送的IP分组随后被GGSN和SGSN各自在相应的位置做上标记。

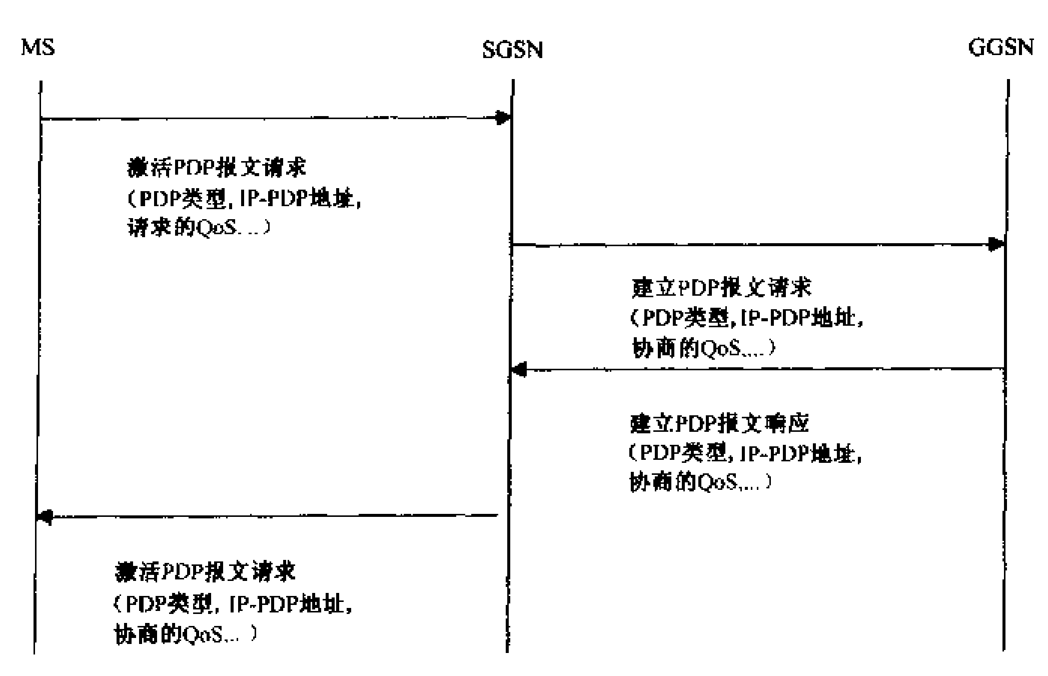


图10.13　PDP报文激活(引自Bettstetter等,1999:Fig.4)

GPRS空中接口

在介绍了GPRS核心网络结构和协议层之后，我们来讨论一下关于GPRS空中接口的更多细节。从RLC/MAC-PDU(参见图10.12)开始，我们将介绍它们是如何映射到GSM时隙——帧结构的。本章的开头讲到了GPRS分组与GSM电路交换呼叫共享GSM帧结构的情况。回忆一下第8章，GSM帧长为4.615 ms，每帧分成8个时隙，每个时隙有114个比特。在GPRS分组结构中，RLC/MAC-PDU经过适当编码和错误检测纠正之后映射到4个时隙的数据部分，每个都来自4个连续帧，称之为无线模块。编码过的RLC/MAC-PDU因此是456比特长。这个无线

模块代表 GPRS 分组数据信道的基本传输单元，GPRS 逻辑业务和控制信道都映射到这个 GPRS 分组数据信道（PDCH）中。在以前就讨论过，PDCH 与 GSM 电路交换业务信道共享时隙、GSM 物理资源。因为每帧有 8 个时隙，根据需求和 QoS 的要求，GPRS 用户最多有 8 个 PDCH 可以利用。这是"按需分配"，而在 GSM 业务信道中，PDCH 只用来传输各自的分组并且在传输结束的时候被释放。此外，PDCH 还可以根据 52 帧组成的多帧结构定义，因此有 240 ms 长。在 GSM 中，每隔 13 个帧就有一个 PDCH 未被使用，因此，每 240 ms 多帧就有 12 个无线模块可以利用，平均每 20 ms 传输一个无线模块。

现在回到用无线模块传输 RLC/MAC-PDU。图 10.14 显示了 RLC/MAC-PDU 的基本结构、编码优先顺序和携带用户数据。RLC 数据部分或者有效载荷的前面加上一个 3 字节的 RLC 头，有效载荷也许是由 LLC-PDU 的段组成的。用携带信号或者控制信息的 RLC/MAC-PDU 替代 RLC 数据和头传输这个信息（Cai 和 Goodman，1997）。一个 MAC 头包括功率控制域（PC）、一个块类型显示器 T 和 3 比特上行链接状态标志 USF。USF 域指定分配给特殊无线模块的移动台（Cai 和 Goodman，1997；Kalden 等，2000）。图 10.15 说明了在 456 比特的无线模块上 PDU 是如何被处理的（间隔是存在的，但是图上并没有标出）。有 4 种编码方案（CS-1 到 CS-4）定义处理类型，处理类型允许空中传输条件的变化。CS-1 提供最强的编码保护，CS-4 的编码保护最小。编码方案 CS-1 到 CS-3 都是用限制长度 4、1/2 码率的卷积编码来做错误纠正的，如图 10.15 所示；CS-4 使用非卷积编码。方案 1 增加40 比特的块检查序列编码（BCS），用来检查卷积码所不能纠正的错误；其他方案使用 16 比特来做 BCS 编码。方案 CS-2 到方案 CS-4 通过增加 USF 编码比特位来保护 USF 域，如图 10.15 所示。表 10.6 比较了每个方案的特点，在每个编码方案中，每个 PDU 在处理后都需要 456 比特，如图 10.15 所示。

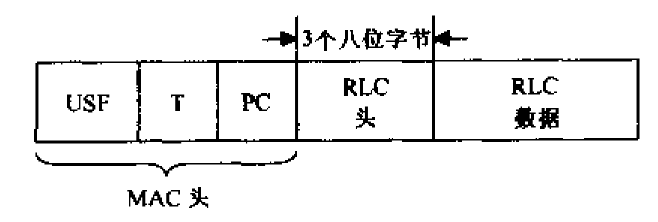


图 10.14 RLC/MAC 协议数据单元

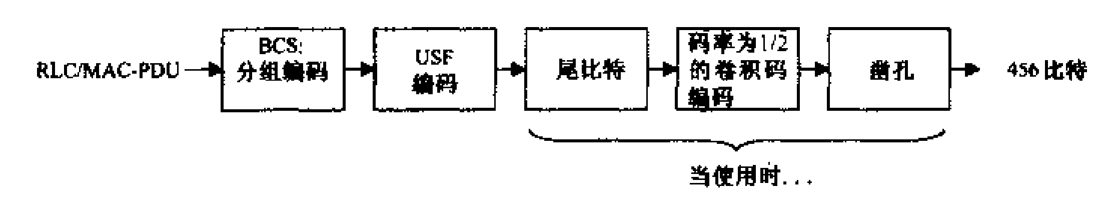


图 10.15 RLD/MAC-PDU 的过程

表 10.6 GPRS 编码方案比较

方案	PDU大小（比特）	BCS（比特）	附加USF编码比特	尾比特	卷积码编码器（比特）输入	卷积码编码器（比特）输出	删除的比特	有效率
CS-1	184	40	0	4	228	456	0	0.5
CS-2	271	16	3	4	294	588	132	0.64
CS-3	315	16	3	4	338	676	220	0.74
CS-4	431	16	9	–	–	456	0	1

表中第 1 列 RLC/MAC-PDU 的大小是在处理前 PDU 允许的大小，用来获得处理后所需的 456 比特。这个 PDU 大小包括 3 比特 USF，与图 10.14 是一致的。CS-4 中 431 字节的最大 PDU 大小与前面所引用的最大 RLC/MAC-PDU 大小的 54 字节是一致的。因此，PDU 大小的范围是 23 字节到 54 字节（包括头）。表 10.6 的第 3 列是 USF 编码所需要的额外的比特。更通常地，所引用的图表明 USF 编码盒是一个“编码前 USF”盒（Cai 和 Goodman，1997；Bettstetter，1999）。除了 CS-4，方案 CS-1 到方案 CS-3 都使用了 1/2 码率的卷积编码。当需要时，可以在末尾加上编码。方案 CS-2 和 CS-3 的卷积编码输出都多于 456 个比特，这些多余的比特随后都被“刺破”或者删除了。表的最后一行是卷积编码输入与 456 比特输出之比，是有效编码率。

表 10.7 GPRS 吞吐率（单时隙每帧；GSM 模块 w/o EDGE）

方案	无线分组数据速率 (Kbps)	LLC 速率 (Kbps)
CS-1	9.05	8
CS-2	13.4	12
CS-3	15.6	14.4
CS-4	21.4	20

因为传输每个大小为 456 比特的无线模块大概需要 20 ms，方案 CS-4 需要传输一个更大的 PDU，提供的吞吐率需比 CS-1 高。从表 10.6 可以清楚地看出，随着方案编号的增加，吞吐率越来越大。表 10.7 显示了在 RLC/MAC 层（无线模块数据速率）和 LLC 层 4 种方案的有效数据吞吐率。体现无线模块吞吐率的速率基于前面表 10.6 中第 1 列提到的无线模块有效载荷再减 3 比特的 USF（Cai 和 Coodman，1997）；LLC 吞吐率是在减去 MAC 头和来自 RLC/MAC-PDU 的 24 字节 RLC 头（Bettstetter 等，1999）后四舍五入得到的。注意，在这里假设传输是基于每帧中的一个时隙作为一个单元的。PDU 最多每帧可以使用 8 个时隙，以增加用户的吞吐率。用来增加 GSM 传输率的 EDGE 技术也能增加这些指数，我们将在下面的章节中专门来讨论 EDGE 技术。

GPRS 逻辑信道携带的用户数据业务和控制、信号信息都映射到分组数据信道 PDCH 中。这些逻辑信道和第二代移动通信中使用的很相似，和其他第三代移动通信也很类似。在讨论 GPRS 逻辑信道，介绍它们如何在 GPRS 系统中使用之前，我们有必要简单介绍一下这些逻辑信道与 PDCH 和无线模块是如何关联的。用户数据是由高层产生的，接着传给 LLC-PDU，然后分段到 RLC/MAC-PDU 中，如图 10.11 和图 10.12 所示。这些 PDU 服务与用来携带用户业务的逻辑信道称为分组数据业务信道（PDTCH）。这些逻辑信道定义成上行链接和下行链接两个方向。

在图 10.11 中，空中接口的控制在 RLC 和 MAC 层进行，空中接口功能的信号和控制信息则通过 RLC/MAC-PDU 传递。这些信号和空中接口功能包括移动台注册、下行链接广播信息（包括基站身份、频率操作和同步信号）、接入请求和接入许可以及寻呼通知，等等。RLC 和 MAC 层为 GPRS 定义了一系列逻辑信道，用来传输这些信号/控制信息。这些消息接着就转化成 RLC/MAC 控制 PDU。用来传递数据业务和控制信息的 GPRS 逻辑信道都列在表 10.8 中，它们以种类和功能分组（Cai 和 Goodman，1997；Bettstetter 等，1999）。注意，这与前面讨论的其他蜂窝系统的逻辑信道很相似，特别注意与第 8 章中 GSM 逻辑信道的相似性。

表 10.8 中的分组广播控制信道（PBCCH）是基站用来给蜂窝内的所有 GPRS 移动终端发送系统信息的，这和 GSM 中的广播控制信道（BCCH）类似。如果 PBCCH 在蜂窝中无法使用，

GPRS 移动终端将接听 BCCH 来获取必要的系统信息,因此电路交换和分组交换的逻辑信道的相互协作是很重要的(Bettstetter 等,1999)。分组公用控制信道和 GSM 控制信道也很相似。与 GSM 及其他蜂窝系统一样,移动台使用随机接入信道(PRACH)告知基站它需要传输分组数据,因此请求基站能够分配给它更多的分组数据业务信道(PDTCH)。基站使用接入允许信道(PAGCH)回答移动台,常规 GSM 的随机接入信道(RACH)和接入允许信道(AGCH)也可以被使用。时隙 Aloha 类型随机接入技术再一次被使用在接入请求中,相关的控制信道 PACCH 专门分配给一个特定的移动台,PACCH 和用在该移动台中的任意方向的 PDTCH 相联系,它用来传输分组数据确认、功率控制信息以及资源分配消息。与 GSM 中一样,基站在等待下行连接分组传输初始化之前用寻呼信道 PPCH 来寻呼移动台。GPRS 也支持多点传送分组给各个移动台。多点传送通知信道 PNCH 就是被用来在传输直接的组消息之前通知各个组成员的。图 10.16 描述了一个移动台寻呼可能的场景(Bettstetter 等,1999)。基站使用 PPCH(或 PCH)发送寻呼请求消息;移动台使用 PRACH(或 RACH)响应信道接入请求消息;基站随后使用 PAGCH(或者 AGCH)发送分配的资源;移动台用 PACCH 确认寻呼。

表 10.8　GPRS 逻辑信道

分类	信道名称	功能	方向
分组数据业务信道	PDTCH	数据业务	双向
分组专用控制信道	PACCH	随路控制	双向
	PTCCH	定时提前控制	双向
分组广播控制信道	PBCCH	广播控制	下行
分组公用控制信道	PRACH	随机接入	上行
	PAGCH	接入允许	下行
	PPCH	寻呼请求	下行
	PNCH	多点传送通知	下行

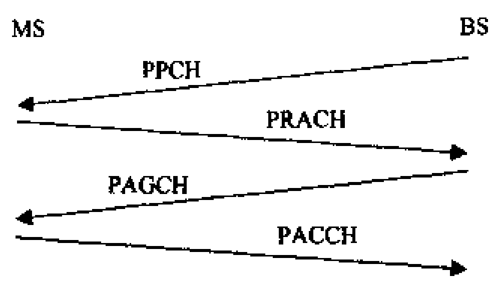


图 10.16　寻呼中的逻辑信道

EDGE/EGPRS

我们修改了 GSM 和 GPRS 空中协议以增强空中接口的传输率,下面就来谈谈这个。有两个基本的改变:在 GSM 中用 8-PSK 调制来替代 GMSK,同时修改无线连接控制模块来改进链路的传输质量(调制技术请看第 5 章)。增强的接口称为 EDGE,用来增强数据率。EDGE 也用在 IS-136 中来增强性能。既然第三代移动通信不能用一种统一的标准来实现,那就用尽量少的标准来实现。有许多论文都是讨论在 IS-136 以及 GSM 系统中执行 EDGE 的(IEEE,1999)。在这里,我们着重讨论 GSM/GPRS(在前面的讨论中,EDGE 中的 G 代表的是 GSM/GPRS。为了与其他的 3G 系统保持一致,把 G 称为全局(global)更恰当些)。EDGE 在分组交换和电路交换中

都有定义。分组交换的 EDGE 称为 EGPRS;电路交换版本称为 ECSD(Furuskar 等,1999a[①])。

现在转移到 EDGE 的细节问题,首先考虑 EDGE 的调制技术,线性 8-PSK 的引入。线性 8-PSK 是在第 5 章讨论的 QPSK 的扩展,包括等间隔的 8 个相位点,这是 QAM 的特例,如第 5 章中讨论的一样(回忆 cdma2000 中使用的 QPSK、8PSK 和 QAM)。图 10.17 描绘了 8 个空间相位位置。正如第 5 章所述,用多相调制能够在给定的带宽内增加比特传输效率。每个传输对应一个符号,一个符号有多个比特信息。在 8-PSK 中,存储了 3 个连续比特,每 3 个比特对应 8 个相位中的一个相位点,3 个比特的间隔就是符号间隔。因此,比特率能够提高 3 倍,这在以前 cdma2000 的讨论中已经提到过。这种调制的缺点是,随着调制相位数目的增加(这里是 8 点),错误率也会相应的增加(第 5 章中,如果传输功率固定,QAM 点之间的最大距离是不能增加的。因此,当点数增加时,它们将靠得更近)。当在衰落信道中以及来自其它移动台或者基站的干扰会导致采用 8-PSK 调制时,将会使包错误率增加。因此,我们有必要改进纠错过程和采用自适应传输速率的方法。下面我们将看到,在 EDGE 中这两种方法都有。

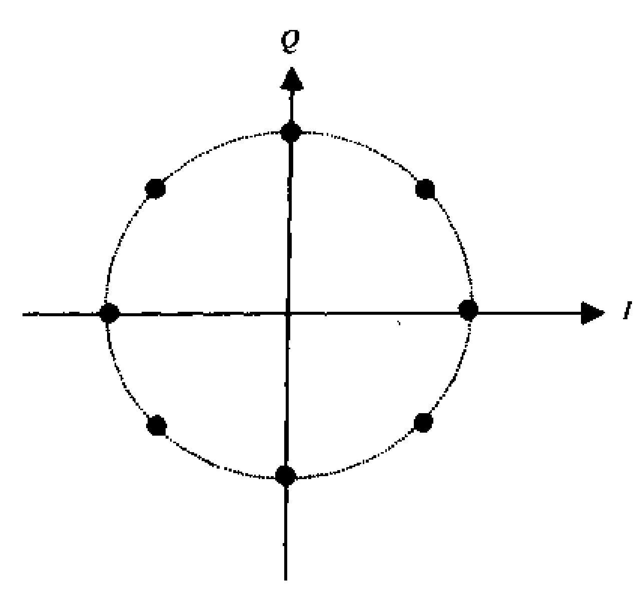


图 10.17 8-PSK

现在讨论在 GPRS 中可能使用的 EDGE 技术(Furuskar 等,1999a)。我们再次注意到,要将效率因子提高到 3,可能的方法是采用 8-PSK。为了使 EDGE 无线接口参数和 GSM 尽量保持一致,采用的频谱形状是线性 GMSK,基本时隙也是与 GSM 基本一样的。略微不同的是,GSM 中的时隙持续时间是 0.577 ms,携带 114 个数据比特,一个用户只能使用一帧 8 个时隙中的一个,每 13 个帧中有 12 个携带用户数据,因此每个用户的基本数据率是 $(12/13) \times 1/8 \times 114/0.577 = 22.8$ Kbps。图 10.18 是 EDGE 的时隙格式,注意时隙中每个域是"符号"而不是"比特"。每个符号间隔内传送的是 8 个 8-PSK 信号中的一个,每个符号携带 3 比特。其次,每个数据域有 58 个符号,因此每个时隙含两个数据域,就有 116 个数据符号,而 GSM 中每个时隙只有 114 比特。比较图 10.18 和 GSM 中的图 8.2,GSM 中的两个 F 比特被数据域代替了(F 比特在 GSM 中是用来表示是否有数据或者控制业务出现的;EDGE 中没有这种差别。回忆 GPRS 中的 RLC/MAC-PDU

① Furuskar, A. *et al*. 1999a. "EDGE: enhanced data rates for GSM and TDMA/136 evolution," *IEEE Personal Communications*, 6, 3(June), 56 – 66.

携带它自己的控制或者数据 PDU),因此一个 EDGE 时隙的分组数据率是(12/13) × 1/8 × 116/0.577 = 23.2 k 字符/s 或者是 69.6 Kbps。如果分组数据用户使用全部的 8 个时隙传输,使用 EDGE 的最大传输率将是这个值的 8 倍或者是 557 Kbps(注意,每个载波在给定 200 kHz 带宽的时候,每个载波的实际传输率是 271 K 符号/s,这与 GSM 中 8-PSK 的使用是一样的,除了传输的是符号而不是比特之外)。由于衰减、传播的不规则以及蜂窝系统中其他不利因素,这个最大值并不一定能达到。由前面讲到的 GPRS 定义的 QoS 包括最大吞吐率和平均吞吐率,EDGE 同样具有这些属性。特别地,当用户以 100km/h 移动时,EDGE/EGPRS 的最大比特率是 384 Kbps,尽管理论上的传输率是 557 Kbps。这与 WCDMA 的最大传输率是一样的。当移动台以 250 km/h 运动时,最大速率降到 144 Kbps(Furuskar 等,1999a)。

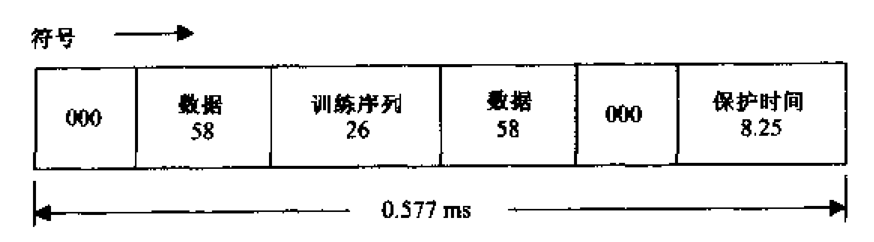


图 10.18 EDGE 时隙格式

在介绍 EDGE 和 EGPRS 时,曾注意到为了使用 8-PSK 调制,无线链接控制有所变化。与 WCDMA 及 cdma2000 一样,比特率的变化是与链路质量所提供的期望变化联系在一起的。回忆表 10.6,在基于 GMSK 的 GPRS 中,无线接口提供了 4 种编码方案 CS-1 ~ CS-4,用来随着链路质量的改变调节传输率。EDGE/EGPRS 又增加了 6 个额外编码方案 PCS-1 ~ PCS-6(Furuskar 等,1999a)。每种方案的主要参数,在表 10.9 中以每个时隙的无线接口比特率为序排列。1/3 码率卷积编码用在新的编码方案中,所有的都是用 8-PSK。注意,这些无线接口的比特率比表 10.7 中代表平均无线模块和 LLC 率的吞吐率要高。表 10.9 中的比特率代表空中接口的有效原始比特率。我们可以通过增加分配给用户每帧的时隙来提高这些比特率。表 10.9 中的不同比特率是通过以最高比特率编码方案 PCS-6 开始的,然后每块连续传输多个比特,接着从 CS-3 到 CS-1。PCS-1 ~ PCS-6 使用 1/3 编码卷积码的性能优于 GMSK 编码方案 CS-1 ~ CS-3 中的 1/2 码率。

表 10.9 EGPRS 中编码方案(基于 Furuskar 等著,1999a,Table 1)

方案	相对码率	调制	无线接口速率/时隙(Kbps)
CS-1	0.5	GMSK	11.4
CS-2	0.64	GMSK	14.5
CS-3	0.74	GMSK	16.9
CS-4	1	GMSK	22.8
PCS-1	0.33	8-PSK	22.8
PCS-2	0.49	8-PSK	34.3
PCS-3	0.59	8-PSK	41.3
PCS-4	0.74	8-PSK	51.6
PCS-5	0.82	8-PSK	57.4
PCS-6	1	8-PSK	69.6

表10.9中10种方案的具体选择取决于无线连接质量条件。EGPRS使用链路控制来适应信道条件,有两种链路控制方法:链路适应和增加冗余(Furuskar等,1999a)。在链路适应中,首先估计链路质量,然后选择合适的编码方案来最大化数据传输。在增加冗余的方案中,最高可能编码首先被选择,用来传输无线模块。如果块传输不成功,就用低一点的传输率来重传块(使用第7章中讲到的ARQ重传)。传输率因此自动适应,并且逐渐转向低数字的编码方案。

EDGE/EGPRS在改进QoS参数,例如吞吐量和分组延时方面,效果如何呢? Furuskar等(1999a)使用了一个常规的大尺度蜂窝层做了仿真。仿真着重于下行链接到Web类型的业务模型,同时考虑了传输距离衰减和对数正态衰落。Furuskar等(1999b)[①]同时做了相关的仿真,使用瑞利分布多径衰落,虽然分组比特率稍微低了点,但是仿真结果还是很相似的。分组比特率低的原因可能是包含了多径衰落的缘故。在这里简单地总结一下Furuskar的报告。尽管报告中也有IS-136(TDMA/136)中的EDGE仿真,但是为了与本节一致,只讲EGPRS结果。

仿真中选择了3部分宏蜂窝。负载随着每个部分用户数量的增加而增加。衡量负载的两个基本参数是:标准延迟参数和频谱效率。后者是每个站点可利用带宽中的比特率,以bps/Hz每站点来度量。这个参数显然和它的名字相符,是用来测量可利用的带宽效率的。标准延时定义成绝对延时,包括队列延时再加上传输延时(单位是s)除以分组长度(Kbps)(更长的分组会明显经历更长的延时,因此必须通过标准化来降低性能对分组长度的依赖)。标准延时的倒数是吞吐率(Kbps)。它代表系统实际的吞吐率,这与前面讲的吞吐率(如无线接口比特特率、无线模块数据率、LLC数据率)略微不同。在讨论IS-95时,也碰到过实际吞吐率或者有效吞吐率这些概念。这些吞吐率和频谱效率模拟结果的总结请看表10.10。表中列出的是90%标准延时,也就是有10%的用户的吞吐率会低于表中的这些值。

表10.10 EDGE/EGPRS与GSM(90%延迟结果)模拟比较(基于Furuskar等,1999a)

载荷	EDGE/EGPRS		GSM		EDGE/GSM(比率)	
(用户数/扇区)	流量 (Kbps)	频谱效率 (bps/Hz/site)	流量 (Kbps)	频谱效率 (bps/Hz/site)	流量	频谱效率
30	25	0.14	14.3	0.03	1.7	4.7
40	21.2	0.19	13.3	0.07	1.7	2.7
50	20.2	0.25	12.5	0.11	1.8	2.3
60	16.7	0.28	8.3	0.14	2	2
70	5.9	0.33	3.7	0.17	1.6	1.9

正如所预料的,当载荷增加时吞吐率会下降(等价于标准延时增加)。这个下降一直很缓慢,直到达到70%的用户,此时,点拥塞开始出现,吞吐量会急速下降(延迟迅速上升)。然而,频谱效率继续增加(管道开始装满,传输分组之间的缝隙开始消失,带宽利用率提高。带宽利用率的提高需要与提供给用户的性能(延时、吞吐率网络操作期望的性能)相平衡。采用EDGE所带来的改进很容易从表中的最后两列中看出来,最后两列是EDGE相对于GSM的相对性能,用吞吐率和频谱效率来衡量。特别地,在大部分的载荷中,吞吐率都低于2:1。

① Furuskar, A. *et al*. 1999b. "Capacity evaluation of the EDGE concept for enhanced data rates in GSM and TDMA/136," *Proc. IEEE Vehicular Technology Conference*, VTC99, 1648 – 1652.

对每个用户的平均比特率的分布也进行了仿真,用来测量系统对各个用户的公平性,详细请看 Furuskar 等(1999a)的著作。特别地,当每个部分的载荷达到 60 个用户时,97%用户的平均比特率超过了 18 Kbps(如果使用每帧中的一个时隙)或者 144 Kbps(如果每帧中的所有时隙都被使用);30%用户的平均比特率超过 48 Kbps(如果使用每帧中的一个时隙)或者 384 Kbps(如果每帧中的所有时隙都被使用)。

在 2.5/3G 系统的讨论中,我们最后讨论 EDGE 对改进 GSM/GPRS 系统性能的效力。在第 12 章中,我们将讨论更高速率的 802.11 WLAN,以及小范围内高比特率方案的蓝牙技术。虽然这些系统所能覆盖的范围比蜂窝系统要小,但是相互之间(包括蜂窝系统)能够互联,提供广泛的覆盖,从而使高速率的多媒体业务在哪里都能被应用。下一章有必要讨论多重接入问题:当同时传输电路交换和分组交换时,怎样分配带宽资源?当需要处理语音、视频和数据业务时,如何调度?

习题

10.1 当 CDMA 码片速率是 3.84 Mcps,路径差异是 78 m 或 240 ft 时能够被分辨;当码片速率是 1.2288 Mcps 时,能够分辨的路径差异是 240 m。解释一下这些语句。

10.2 (a) 怎样在不降低扩频增益的情况下增加多码 CDMA 的传输速率?"用在并行中的码字必须相互正交"是什么意思?"接收机的自干扰"是什么意思?正交是如何克服自干扰的?

(b) 请画出多码 CDMA 在基本速率 3 倍情况下的图,说明是如何通过串并转换来达到 3 倍速率的。

(c) 解释为什么多码 CDMA 的包络会变得更宽。为什么需要多重瑞克接收机?

10.3 为了使瑞克接收机有效,扩频信号带宽必须比相干带宽大?为什么这样会导致频率选择性衰退?为什么为了解析多径回声,必须有频率选择性衰退?请解释之。

10.4 考虑 WCDMA 系统

(a) 为什么用户数据传输率低于空中接口的 1/2 传输率?为什么局域环境的传输率比广域环境高?

(b) 在上行链接无线信道中,如何获得 5.76 Mbps 的速率?描述在扩频增益是 4 的情况下是如何实现的。

(c) 简单解释一下,为什么 QPSK 用来下行传输,而 PSK 用来上行传输(QPSK 的速率是 PSK 的两倍)?

10.5 在书中讲述了 4 种 QoS 业务类型。为什么在连续信息实体中,实时会话业务需要低延时并且限制时变,而实时流类型只需要限制时变。为什么电子邮件业务对投递延时相对不敏感?并以你自己生活中的非实时交互业务为例,同时阐述它们的突发特性。

10.6 图 10.1 显示了两个 WCDMA 物理信道并行传输。解释两个信道怎样能做到同步传输,以及在接收端能够分别被检测。

10.7 考虑 WCDMA RACH 信道

(a) 为什么传输率是 120 Kbps 时扩频增益系数是 32?为什么在考虑功率后需要的扩频增益是 128。在这种情况下,为什么通过 RACH 信道的分组数据少于 150 比特?

(b) 画出接入过程图,同时包括移动台和基站之间的各种消息传输,需要重传 RACH 前导码多次的情况。

10.8 (a) 比较 IS-95B 和第 8 章中的 IS-95。

(b) 为什么使用 Walsh 编码的多码操作能够增加比特率?

(c) 考虑 IS-95B 分组数据传输的两个例子,传输率均是 115.2 Kbps。1200 比特分组的有效传输率是 96 Kbps,而 800 比特分组的有效传输率是 80 Kbps,解释它们的有效传输率为什么不同? 如果考虑反向接入过程,有效传输率将更低,为什么?

10.9 在 cdma2000 中

(a) 表 10.2 介绍了各种反向信道无线配置。配置 3 中,为什么在比特率增加时传输增益反而会下降。在高比特率系统中,你期望用哪一种系统;在低比特率系统中,你期望用哪一种系统?

(b) 图 10.3 的 307.2 Kbps 是如何获得的? 在输入 R-SCH 结构时,实际比特率是多少?

10.10 比较 cdma2000 和 IS-95 的各种信道,包括传输的两个方向。它们哪些是相同的,哪些是不同的? 绘出在每个方向上的各种信道图,在图中给相似的信道分组。在这个图中所显示的与 IS-95 向后的兼容性如何?

10.11 比较 WCDMA 和 cdma2000 的逻辑信道,包括上行和下行。

10.12 正文中介绍了 cdma2000 移动功率增加和注册的过程。画出这个过程的图,同时标明正向和反向传输的各种信号和消息,直到移动台进入空闲状态。

10.13 表 10.4 是 cdma2000 1 × EV – DV 系统中前向信道的各种分组数据传输率。

(a) 用你自己的话解释各种数据率是如何获得的。找出图 10.7(A 点和 B 点)中各种情况下输出模块的符号率。图 10.7 的信道间隔的作用是什么?

(b) 与传输高比特率相比,移动台在使用 8-PSK 和 16-QAM 接收 38.4 Kbps 前向传输速率时的 SIR 值要更小,这是为什么?

10.14 图 10.11 中的 GPRS 层结构应用于两种场景。第 1 种,移动台借助 GPRS 连接因特网,并通过因特网连到相应的 Web 地址。假设移动台是图 10.8 中的 M1,因特网是 PDN。假设 Web 地址需要两跳才能连接到 GGSN。结合图 10.8、图 10.9 和图 10.11,解释一个应用数据 A-PDU 是如何从 M1 出发到达 Web 地址的。第 2 种应用情况是,M1 传输分组数据到 M3。PDN 再一次由两个路由器来表示,但是只是从一个 GCSN 传输数据到另一个 GGSN。请再一次结合各种相关的图,解释 M1 中的数据是如何到达 M3 的。请表明各系统中各个层协议是如何被利用的,以及在各个层中头信息是如何被加上和去掉的。为了简单起见,假设 A-PDU 很小,不用分段。

10.15 (a) 比较表 10.8 中的 GPRS 逻辑信道和第 8 章中 GSM 的逻辑信道。

(b) 比较 GPRS 和 GSM 的寻呼过程。

10.16 (a) 如果每帧使用一个时隙,GPRS 编码方案 CS-4 的有效吞吐率是 22.8 Kbps。有效吞吐率的最大值是什么(又称为无线接口率,见表 10.9)?

(b) CS-1 到 CS-3 的有效吞吐率在表 10.9 中给出,比较表 10.6 中最后一列的相对有效率。

(c) 验证表 10.7 中的各种 GPRS 吞吐率。

10.17 用你自己的话解释,表 10.9 的各种比特率是如何获得的?

10.18 比较表 10.9 和表 10.10 的吞吐率。为什么表 10.10 中的比表 10.9 中的最大值 69.6 Kbps 低得多? 为什么表 10.10 中吞吐率的值随着载荷的增加而降低?

第11章 蜂窝系统接入和调度技术

从以前各章的讨论中可以看到,在研究蜂窝系统时,移动台和基站间进行通信的无线信道限制了通信能力。尤其是当使用TDMA或CDMA作为多址接入技术时,只有有限数量的用户适于在指定的频段通信。这样,用户在传输时必须依据具体情况得到许可,才能使用指定的频段、时隙或码字。为了得到传输许可,用户要在接入时提出接入请求。由于只有有限个接入信道,比如有限个频率和时隙或有限个频率和码字,造成了同时接入的用户之间的干扰或者称为"冲突"。通常情况下,在第8章和第10章关于第二代和第三代蜂窝系统的讨论中,使用基于时隙Aloha(slotted-Aloha)技术的随机接入方法和避免冲突的机制进行接入。在电路交换的语音呼叫系统中,一旦接入请求被响应,用户就会根据具体情况在给定的频率上分配一个时隙或码字,并且该时隙或码字一直持续到呼叫结束。在第10章所述的第三代系统中,基于分组交换的数据传输则采用了更为复杂的机制。

如第10章所述,第三代系统要支持各种具有不同流量特征的基于分组的数据业务。这些业务可以被分为以下4类:实时的语音和数据流业务;非实时的交互和突发业务。这里为了便于讨论接入问题,将以上各种分组交换数据业务分为两大类,以简化讨论,它们是时间连续业务和突发业务。时间连续业务以语音和视频为代表。这类业务一旦成功地接入系统,就必须在一定时间内以一个固定速率进行传输,就像电路交换那样,需要给它们分配一个确定的无线信道或网络资源。所以,一旦接入成功,系统即需提供能满足持续一段时间的数据率的容量。而在突发业务中,每个突发包含一个或一串包,它需要多个接入请求。根据系统设计的不同,可以每次许可接入一个包,也可以每次许可接入一个突发的一部分包,或者许可接入一个突发的所有包。每个突发至少需要一次接入请求。而且,在每个允许接入的间隔内,无线信道容量或传输速率被分配给一个固定的用户,这个用户可以是上行或下行,被分配的传输速率可以有所不同,详见第10章。在每个间隔中具体分配多少资源取决于待传输的数据特征、用户的QoS特征和业务类型。基站必须采取某种分组调度机制在上下行方向上实施各帧之间的资源分配,以保证用户间的公平,这种分配可以依据业务类型、QoS特点和信道状况来进行。

由此看出,接入被允许后,分组业务的传输还需要适当地接入控制和传输调度。所有这些问题都将在本章讨论。值得注意的是,接入控制和调度策略的操作都被定义在MAC层或数据链路层中的媒介接入子层(见第10章)。我们还将在下一章看到,在无线局域网WLAN和个域网中,这些操作被定义在MAC子层或位于其下面的物理层。那么,一个首要的问题是,什么是最适当的接入机制?怎样比较不同的接入机制?无论是电路交换还是分组交换,这个问题人们都十分关注。一旦接入被允许,如上所述,数据分组业务就需依据业务类型、QoS要求、信道状况进行适当的调度。那么,怎样才是合适的调度机制?所有这些设计问题都将在下面详细地讨论。

在11.1节,从应用最广的时隙Aloha接入机制入手,讨论接入控制。简要分析它的性能,先是在没有衰落的环境下,然后在瑞利(快)衰落环境下讨论。接着,提出一些改进随机接入性能的方法,并将改进后的性能与原始的性能进行比较。比如,在第10章已经提到的在WCDMA

接入过程中采用的 CPCH。在 11.2 节,我们提出了一种能够将不同种类的业务结合起来,并在 TDMA 系统中传输的机制。最常见的是将语音和突发数据结合起来传输。这种机制涵盖了接入控制和调度两个方面。先介绍较早出现的无线分组多址接入(PRMA),然后介绍诸如 PRMA++ 和与其相关的 DQRMA 等改进技术。在 11.3 节,通过讨论一系列适用于 TDMA 和 CDMA 系统的调度算法来结束本章的内容。其中的一些算法是为未来的全分组第 4 代蜂窝系统提出的。在后面的介绍中我们可以看到,这些调度算法中的很多都是从有线分组网的调度算法中演变而来的。

11.1 时隙 Aloha 接入

如上所述,时隙 Aloha 接入(Slotted-Aloha access)技术和它的变形已经广泛地应用于第二代和第三代蜂窝系统中。我们已经在第 8 章和第 10 章分别给出了该技术在第二代和第三代蜂窝系统中的应用,这里将给出技术的细节、该技术的一个简单的性能评价、衰落信道对于该技术的影响,以及为提高性能对此技术做的一些改进。

时隙 Aloha 接入是 Aloha 接入中的一种。Aloha 接入是由夏威夷大学的 Norman Abramson 和他的同事们在 20 世纪 70 年代早期(Abramson,1973)[①]提出的。Aloha 接入机制的基本思想非常简单:想传就传。如果有冲突,也就是有两个或更多的用户在重叠的时间内都要发送数据,那么用户就在一个随机时间间隔后重复发送(随机的重传间隔可以减少重传时的冲突概率)。这种最简单的随机接入方式称为纯 Aloha。尽管这种方式很显然有冲突的可能,但有证据表明,这种技术在业务为突发模式且包长较短的情况下,能提供可观的吞吐量。这是种可行的方案,特别是它非常简单易行。全球流行的以太网使用了一种 Aloha 接入的改进形式,它能提供更高的吞吐量,而且仍然十分简单易行。关于以太网技术的简要介绍将在下一章给出。在下一章我们将会介绍另外一种系统无线局域网,它采用的也是 Aloha 技术衍生出的一种接入技术。

现在讨论时隙 Aloha 接入技术。该方法本质上同纯 Aloha 技术相同,惟一的不同在于该方法划分了时隙,被接入的分组必须以时隙为基本单位进行传输,而且必须在各个时隙的起始处开始传输。时隙结构见图 11.1。图中的时隙为 T s 长,而且周期性地重复。一个被接入的分组在其中一个时隙传输,在另一个时隙则有两个冲突的分组欲同时传输。这里,接入过程同样是“想传就传”,只是必须在一个时隙内。如果一个用户检测到冲突,比如,多于一个的用户企图在同一个时隙进行传输,那么该用户在一个随机数目的时隙间隔后再发出接入请求,如图 11.2 中的例子所示。首先是一个成功的接入请求(用户被告知接入成功的时间没有显示在内),随后,图中显示了一个冲突,冲突了的用户被告知的时间也显示在其中。然后是到重传之间的间隔,用户随机选取一个时隙发送重传请求。在所考虑的蜂窝系统中,基站会告知移动台成功接入的消息。如果用户没有从基站处得到成功接入的确认,就说明发生了冲突或者接入不成功。第 8 章和第 10 章所讨论的各种蜂窝系统所采用的时隙 Aloha 接入方法都是基于这种机制的,只在个别处有些改进。这些改进包括重传时增加传输功率,或以时隙为单位增加重传间隔。下一节将详细介绍以概率 p 为标志的另一种改进,重传请求为概率 p。另一种改进是,

① Abramson, N. 1973. "The Aloha system," Ch. 14, in *Computer Networks*, ed. N. Abramson and F. Kuo, Englewood Cliffs, NJ, Prentice-Hall.

一帧内只允许一次重传，一帧包含一定数目的时隙，与前面几章系统中帧的概念相同。另外，在有的改进中，只有特定的时隙可用做接入请求。

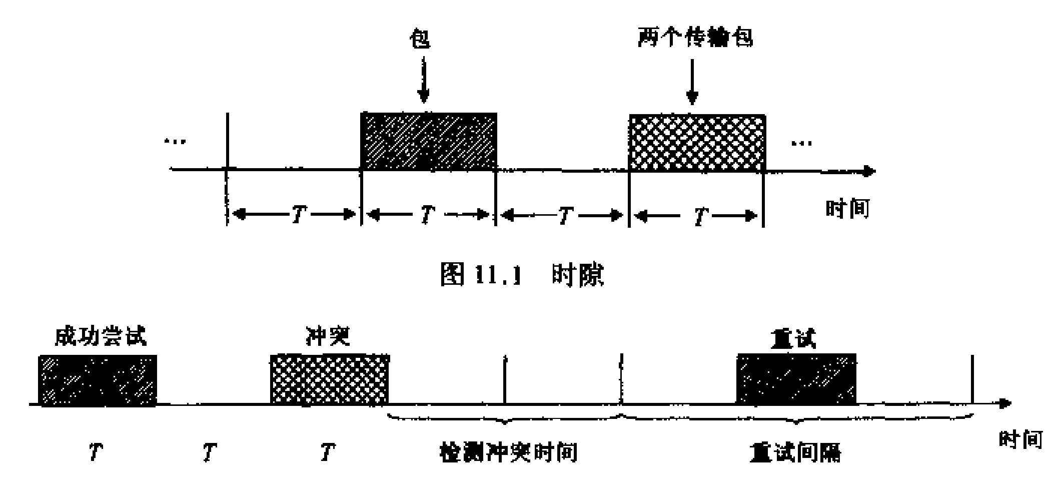


图 11.1 时隙

图 11.2 时隙 Aloha 的过程

现在只考虑如图 11.1 所示的最基本的时隙 Aloha。注意到，系统最大吞吐量或容量，是每 T s 一个包。例如，如果时隙长度为 5 ms，每秒最多传输 200 个包。系统容量，即在给定频率处一个无线接入信道的最大允许吞吐量。比如，所有使用该接入信道的用户的吞吐量的总和不能超过该值。如果潜在用户的数量增加，那么每个用户能分到的吞吐量就会相应地减少。怎样将时隙 Aloha 的性能与这种以每时隙包的个数为单位的最大吞吐量或容量进行比较呢？通过一个简单的分析，我们将看到，最大可能的时隙 Aloha 的吞吐量为 e^{-1}，即 0.368，或者 $0.368/T$ 包/s。事实上，一个系统必须远远低于这个最大值进行工作。时隙 Aloha，与别的随机接入系统一样，在接近最大极限时呈现出不稳定的状态，且必须被控制在一个稳定状态下（参见 Schwartz，1987：p.430）[①]。一个简单的稳定该系统的方法是在不成功的重传请求后增加再次进行随机重传的时间间隔，这是以上改进中的一种。使用时隙 Aloha 会大大降低系统允许的最大吞吐量，但由于操作简单，还是得到了广泛的应用。对于如上所述的例子，如果每个时隙5 ms，最大吞吐量为每秒 0.386 × 200 = 77 个接入。如果系统采用此速率的 1/2 来保证稳定性，那么每秒最多只有 38 个包允许传输。如果每个基站有 100 个终端，那么平均每个终端每 2.5 s 可以得到一次接入的机会。这并不是很难办到的，所以 Aloha 机制才会被广泛地应用。正是在流密度不是很大的情况下，操作才能够很有效并且操作简单。它最先是为解决突发数据业务而设计的，因而对于此种业务尤其是在以太网和无线局域网中的突发数据业务非常有效，很适于解决偶尔发生的接入请求。

为了得到一个简单的描述时隙 Aloha 性能的表达式，使用下面这个通用模型（Schwartz，1987）：假设在一个小区内移动上行接入请求的产生满足泊松分布，平均为每秒 λ 个请求。每个请求对应一个长度为 T s 的接入包，也就是说，在一个 T s 的给定时隙内平均有 λT 个接入请求。泊松分布的这个假设隐含了接入请求在不同时隙之间的随机性和独立性。T s 时间内（一个时隙）发生 K 个接入请求的概率为

$$P(K) = (\lambda T)^K e^{-\lambda T}/K! \tag{11.1}$$

① Achwartz, M. 1987. *Telecommunication Networks: Protocols, Modeling, and Analysis*, Reading, MA, Addison-Wesley.

其中，一些请求会冲突，它们在一个随机时间间隔后重传。这些重传请求显然是与原始请求相关的，尽管这样，还是假设重传请求满足泊松分布。这样的假设主要是为了简化分析过程，但是它也存在其内在的合理性，原因如下：在较长时间后重传将会使被分配的时隙相隔很远从而减少冲突；重传会随机地选择时间间隔；在一定间隔后的某个时隙，传哪个重传请求与具体的用户和此前的事件无关。这样，在给定时隙时，就会既有新产生的随机接入请求，又有以前时隙的重传请求。假设新接入请求和重传请求的总和的平均速率为每秒 λ' 次，其中 $\lambda' > \lambda$。在每 T s 的时隙间隔内有 K 个混合请求的概率如式(11.1)所示，我们将其中的 λ 换为 λ' 即可。

现在介绍几种常用的描述 Aloha 随机接入性能的方法。假设 S 代表 T s 内平均新接入的数目，G 代表 T s 内平均新请求和重传请求的总数目，即 $S \equiv \lambda T$，$G \equiv \lambda' T$。平均每时隙的新请求数目 S 是总的吞吐量的度量(由于这里不考虑拥塞，成功请求的数目就等于新接入请求的数目，即所有的请求最终都会被满足)。式 11.1 中的 λT 可以用 S 替换，同样地，可以用 G 替换 $\lambda' T$，表示 T s 内两种请求总数为 K 的概率。特别地，T s 内没有冲突的概率为 e^{-G}，T s 内至少有一个冲突的概率为 $1-e^{-G}$，平均每 T s 的平均重传次数为 $G(1-e^{-G})$，由于 G 等于平均每 T s 内的总请求，所以有 $G = S + G(1-e^{-G})$，即

$$S = Ge^{-G} \tag{11.2}$$

这个表达式将吞吐量 S、平均每个时隙内成功请求的数目，与一个时隙内总的请求次数联系在一起，而后者又可由时隙 Aloha 的吞吐量公式给出(Schwartz, 1987)。式(11.2)可以有一种不同的更简单的方法导出：在一段很长的时间间隔 T' 内($T' \gg T$)，成功接入的概率为成功接入的次数 ST' 与总的请求次数 GT' 的比，即 ST'/GT'。但这只是没有冲突的概率，即 e^{-G}，令两式相等，我们得到式(11.2)。

式(11.2)的曲线如图 11.3 所示，可以看出，最大吞吐量为 $e^{-1} = 0.368$，与前面所述相同，其最大值在 $G=1$ 处得到。还可看到，当 $G>1$ 时随着 G 的增加吞吐量下降，表明系统进入了不稳定状态，这点在前面的分析中也提到了。显然，用户只可能在小于 S 的情况下操作。有趣的是，纯 Aloha 系统的性能曲线与之极为相似。由于纯 Aloha 没有时隙的概念，用户可以在任意时刻传输而不受时隙边界的限制，因而如果包长为 T，则在 $2T$ 时间内都会发生冲突(请读者自行证明此结论)。如果用泊松分布来分析，用 $2T$ 代替 T，那么纯 Aloha 的性能公式为 $S = Ge^{-2G}$，与式(11.2)相比，吞吐量降为 $e^{-1}/2 = 0.18$，在 $G = 1/2$ 时获得此最大值。时隙的引入使最大吞吐量变为原来的两倍。

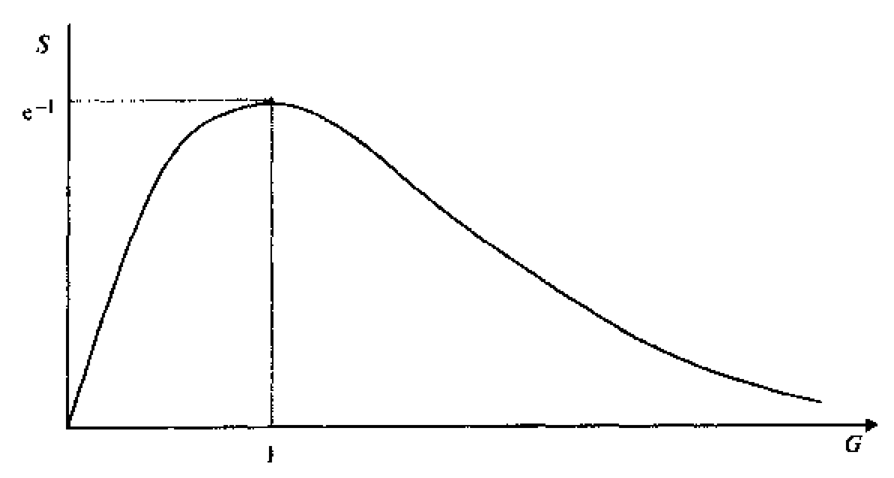


图 11.3 时隙 Aloha 的性能曲线

衰落对于接入控制的影响:"捕获效应"

考虑时隙 Aloha 在接入过程中衰落效应的影响。我们将看到,在信道存在衰落时,系统最大吞吐量比无衰落时有明显增加!这是由于随机接入的一个特性"捕获效应"造成的,该效应早在 Aloha 研究阶段(Abramson,1977)[①]和蜂窝系统性能分析中(Goodman 和 Saleh,1987[②];Yacoub,1993[③])就已经开始被人们所认识了。如果某移动台的信号强度很大,以至于即使与其他移动台同时传输仍然可以被正确地接收到,那么我们就称它"捕获"了信道。比如,距离基站近的移动台与同一小区中那些距离远的相比,就更容易"捕获"。接收功率由于移动台位置、阴影效应、快衰落现象等因素的影响可能会相差若干分贝。这样,蜂窝系统中的功率控制就显得十分重要,尤其是在前些章所提到的 CDMA 系统中。移动台通常有两种功率控制方法,一种是较粗糙的开环功率控制,另一种是较精确的闭环功率控制。在开环控制中,移动台通过估计其所接收到的从下行信道的基站发送的功率,来估计它应该采用的发射功率(Viterbi,1995)[④]。这使得由于路径损耗和阴影衰落造成的移动台功率的差别减小。在闭环控制中,需要基站针对某个具体移动台发送增大或减小功率的消息,来抵消变化更快的瑞利或莱斯衰落。

下面考虑接入机制。移动台可以接收到基站的信号并且估计出它应使用的功率,这些都可在基站检测到该移动台之前完成。这样,移动台就可实现通过调整发射功率来补偿路径损耗和相对变化较慢的阴影衰落的目的。但是,由于它还没有被基站认证(这恰恰是接入过程所要做的事情),闭环控制是不可能的。所以,由于快衰落的影响,基站接收到的不同移动台的功率是不同的,一个处于深衰落中的移动台即使处于较好的位置,也肯定不会被基站"听到",这种在接入过程中基站所接收到的信号功率的差异导致了如上所述的捕获效应的发生。对该效应的分析简述如下:沿用 Arbak 和 Van Blitterwijk 的方法,假定为瑞利衰落(Arbak 和 Van Blitterwijk,1987)[⑤]。特别地,假定一种均匀的情况,基站从各移动台接收到的平均功率都是 P_0。如上所述,这种假设意味着各移动台的开环功率控制要操作得很准确。还假设基站接收到的各移动台的信号差别是充分随机的(当然,没有任何理由说这些不同的信号间有相位同步)。需要格外注意这些接收信号功率,因为它们会在基站处叠加。

考虑第 i 个接收信号,回忆第 2 章内容,在瑞利衰落下,接收信号功率服从如下均值为 P_0 的指数概率密度函数

$$f_{P_i}(P_i)=\frac{1}{P_0}\mathrm{e}^{-P_i/P_0} \tag{11.3}$$

假设,在某个时隙,有 n 个相干扰的包被接收。将这些功率相加得到随机变量 T_n

$$T_n=\sum_{i=1}^{n}P_i \tag{11.4}$$

---

① Abramson, N. 1977. "The throughput of packet broadcasting channels," *IEEE Transactions on Communications*, COM-25, 1 (January), 117 - 128.

② Goodman, D. J. and A. A. M. Saleh. 1987. "The near-far effect in local Aloha radio communications," *IEEE Transactions on Vehicular Technology*, VT-36, 1 (February).

③ Yacoub, M. D. 1993. *Foundations of Mobile Radio Engineering*, Boca Raton, FL, CRC Press.

④ Viterbi, A. J. 1995. *CDMA, Principles of Spread Spectrum Communication*, Reading, MA, Addison-Wesley.

⑤ Arbak, J. C. and W. Van Blitterwijk. 1987. "Capacity of slotted Aloha in Rayleigh fading channels," *IEEE Journal on Selected Areas in Communications*, SAC-5, 2(February), 261 - 269.

以 nP_0 为均值的随机变量 T_n 的概率密度函数服从如下所示的 Gamma 分布

$$f_{T_n}(T_n) = \frac{1}{P_0}\frac{(T_n/P_0)^{n-1}}{(n-1)!}e^{-T_n/P_0} \tag{11.5}$$

假设在相同的时隙内有一个包欲以功率 P_S 接入，这与 n 个干扰接入包一样。同样假设功率 P_S 满足均值为 P_0 的指数分布。如果该包的功率大于其他所有 n 个包的总和，那么它将获得传输机会。特别地，当 $P_S/T_n \geqslant z_0$，$z_0 \geqslant 1$ 时，捕获成功；相反，当 $P_S/T_n < z_0$ 时发生冲突。其中，z_0 通常称为捕获率，取决于系统参数。在计算时通常取为 2，也就是 3 dB。

对于此测试包的成功捕获意味着传输成功，在前一节关于时隙 Aloha 的讨论中知道成功传输概率为 S/G，所以，若考虑瑞利衰落信道，对捕获成功率的计算可得到 S/G 的期望值。若用 $p_c(z_0)$ 表示成功捕获概率，则 $p_c(z_0)$ 的计算过程如下（Arbak 和 Van Blitterwijk，1987）。测试包将会遇到另外 n 个冲突包，包括新的包和重传的包两种，它们都会经历瑞利衰落。在一个时隙中有 n 个包同时被接入的概率如式（11.1）所示，我们将其中的速率参数改为包含所有包。$P(n)$ 的概率由 $G^n e^{-G}/n!$ 给出，G 代表瑞利衰落环境下平均一个时隙接入请求的数目。这个表达式被用来计算捕获概率 $p_c(z_0)$。注意到，捕获概率即为没有冲突的概率，也就是（1－冲突概率），所以有下式

$$p_c(z_0) = 1 - \sum_{n=1}^{\infty} P(n)\,\mathrm{Prob}\left(\frac{P_S}{T_n} < z\right) \tag{11.6}$$

在这里，表达式 $\mathrm{Prob}(P_S/T_n < z_0)$ 定义为当 n 个相干扰的包被传送时在一个时隙内的冲突概率。作为验证，考虑捕获概率 z_0 很大，并接近 ∞ 的情况，$\mathrm{Prob}(P_S/T_n < z_0) = 1$，则式（11.6）简化为 $p_c(z_0) = 1 - \sum_n P(n)$，这正是没有干扰包时的概率，即式（11.1）中使用 G 为泊松参数时的 e^{-G}，这样可得到 $S/G = e^{-G}$，同期望的相同。

更一般的情况是，假设 z_0 为有限值，为确定 $p_c(z_0)$ 必须计算出 $\mathrm{Prob}(P_S/T_n < z_0)$。计算方法如下：使用一个与第 9 章类似的随机变量的二维变换（Arbak 和 Van Blitterwijk，1987）。令 $z_n \equiv P_S/T_n$，$w \equiv T_n$。计算 (P_S, T_n) 到 (z_n, w) 的转移矩阵，结合 w，可以计算 z_n 的概率密度函数，然后可以依此计算捕获概率，得到同第 9 章相同的结果

$$\begin{aligned} f_{z_n,w}(z,w) &= f_{P_S,T_n}(P_S,T_n)\left|\frac{\partial(P_S,T_n)}{\partial(z,w)}\right| \\ &= f_{P_S}(zw)f_{T_n}(w)w \end{aligned} \tag{11.7}$$

用 zw 代替式（11.3）中的 P_i 得到式（11.7）中的第 1 个式子，将式（11.5）中的 Gamma 函数代入可得到第 2 个式子。这样得到

$$f_{z_n,w}(z,w) = \frac{w(w/P_0)^{n-1}}{P_0^2(n-1)!}e^{-(wz+w)/P_0} \tag{11.8}$$

这样，z_n 的概率密度函数就可以通过将式（11.8）与从 0 到 ∞ 的变量 w 的值整合在一起得到。读者可以自行证明最终的概率密度函数如下式所示：

$$f_{z_n}(z) = n(z+1)^{-n-1} \tag{11.9}$$

回忆知道变量 z_n 代表 P_S/T_n，而我们所关心的是计算 $\mathrm{Prob}(P_S/T_n < z_0)$ 的值，它可以在 $z < z_0$

时对式(11.9)积分得到,也就是 z_n 的累积分布函数。经计算该值等于 $1-(z_0+1)^{-n}$。用该值替换式(11.6)中的 $\text{Prob}(P_S/T_n<z_0)$,计算关于 n 的求和,对表达式进行化简,最终得到下式:

$$S=Ge^{-G\frac{z_0}{z_0+1}} \tag{11.10}$$

与式(11.2)中的时隙 Aloha 的性能进行比较,显而易见,由于瑞利衰落造成的捕获效应增加了系统的平均吞吐量。现在 S 的最大值是 $e^{-1}(z_0+1)/z_0$,而不是以前的 e^{-1}。比如,设捕获概率 z_0 为 3 dB 即 2,按照上面的分析,S 的最大值为 0.55,增大了 50%。但是,我们不能稳定地获得此吞吐量。随后,Zorzi 和 Rao 在捕获时隙 Aloha 的稳定性方面的工作显示出,在瑞利衰落情况下,需要进行重传控制来保证时隙 Aloha 接入机制的稳定性(Zorzi 和 Rao,1994)[①]。

他们分析指出,若要保证稳定性,则每个时隙新接入的平均吞吐量 S 必须小于 $2/\pi\sqrt{z_0}$。当 $z_0=2$ 时,能保证接入机制稳定性所允许的最大 S 值为 0.45,与其对应的 G 的最大值为 0.75。如果没有捕获效应,当 $G=0.75$ 时,由式(11.2)可知,$S=0.35$。说明系统位于该点时,由于捕获效应,系统获得了 29%的潜在增益。但是,需要增加每次重传时的重传间隔来保证稳定性。当接入速率降低时,由于捕获效应带来的系统增益会有所减小。GPRS 系统中的捕获效应,同本书讨论的其他系统相同,使用了时隙 Aloha 接入。有仿真表明,这种系统的吞吐量比没有捕获效应的系统大大增加了(Cai 和 Goodman,1997)。本节分析的最本质的不同在于,随机接入中的包排队等待发送。这样,前面提到的不稳定状态就不会出现,图 11.3 中的性能曲线就不会折下来,而是固定在一个饱和值上,但为此付出的代价是延时的增加(Cai 和 Goodman,1997)[②]。

改进的接入策略

为了提高系统吞吐量,人们在如上所述的基本时隙 Aloha 接入技术上进行了一系列改进,形成了一些新的方法,这里简要介绍几种。其中,某些技术已经被应用于第 10 章所介绍的第三代移动通信系统中。

首先,考虑接入过程中的功率提升(power ramping)的概念。该方法应用于第 8 章介绍的第二代移动通信系统。同样,接入中的功率提升也在第 10 章介绍的第三代 CDMA 系统中有所应用。在该方法下,用户首次接入所选用的功率值明显小于通过开环控制估计得到的正常值。注意到由于测量误差和衰落等原因,估计结果往往不够准确。考虑到衰落的可能性,往往系统会预留较大的衰落空隙,这导致了冲突的概率大大增加。相反,如果移动台以一个低于正常值的功率开始请求接入,然后每次重传,功率都会有一定的增加(Olafsson 等,1999[③];Moberg 等,2000[④])。第 10 章中的另一个问题是移动台应该不停地重复发送较短的导频来作为接入请求

---

① Zorzi, M. and R. R. Rao. 1994. "Capture 和 retransmission control in mobile radio," *IEEE Journal on Selected Areas in Communications*, 12, 8(October), 1289 – 1298.

② Cai, J. and D. J. Goodman. 1997. "General packet radio service in GSM," *IEEE Communications Mag.*, 35, 10(October), 122 – 131.

③ Olafsson, H. *et al*. 1999. "Performance evaluation of different random access power ramping proposals for the WCDMA system," IEEE Personal Indoor Mobile Radio Conference.

④ Moberg, J. *et al*. 2000. "Throughput of the WCDMA random access channel," Proc. IST Mobile Communication Summit, Galway, Ireland, October.

直到接入被授权，再发送较长的信息，还是直接发送信息本身，然后不停地重复，直到接入被授权。任何一种情况都应用了功率提升。后者操作延时更长，但是一旦接入被允许，就可保证消息能够被接收到。而前者使用前缀来请求接入，操作时间缩短了，但是，在接入被允许后，消息本身却很有可能由于衰落等原因而丢失（Olafsson 等，1999；Moberg 等，2000）。无论如何，由于较短的前缀的应用，干扰和冲突的概率确实降低了。

作为增大吞吐量的又一新技术，CPCH（上行链路公共分组信道接入机制）曾在本章的引言中被提到过。在第 10 章讨论 WCDMA 时，CPCH 作为 RACH 接入方法的一种补充，曾被简要地介绍。这两者之间的不同在于初始接入阶段。在 RACH 接入中，移动台从基站的广播信道中选择一个 RACH 子信道，依据下行信道测得的接收功率来估计初始发送功率，然后随机选择一个码字或标签来发送一个 1 ms 长的导频。如果基站检测到导频，就在 AICH 信道将该标签返回，然后终端继续发送一个 10 ms 或 20 ms 的 RACH 消息。如果其他的接入终端也使用了同一个标签，并且也对 AICH 消息做出了回应，就会产生冲突。CPCH 机制为了减少由此而产生的冲突，使用了一个不同的随机产生的标签，通过发送一个冲突检测（CD）导频来代替较长的 RACH 消息。基站对这个新标签通过 CD 认证信道（CDICH）的回复进行确认。如果回复信号正确，移动台则启动一次新的传输。正是 CPCH 的短导频和其中传送的随机生成的新标签的结合使得接入性能有所提高。Moberg 等（2000）的仿真结果表明，基站进行接入授权后，通过移动台随机选择一个新的标签的方法可以大大提高接入机制在高载荷下的吞吐量。通过二次随机化可以明显降低多个消息冲突的可能。

11.2 综合接入：语音和数据

以前各节主要涉及时隙 Aloha 在上行信道的随机接入，且只考虑初始接入，而并没有考虑当接入成功后信道资源的合理分配。本节将把接入控制和信道资源的分配结合起来进行讨论。问题的范围还将扩大到包括不同种类的有不同流量特性和不同服务质量要求的业务。就像本章的引言所述，通常满足于考虑两种基本类型的业务：一种是连续时间业务，比如实时语音；一种是突发数据。这里关注它们的接入控制和资源分配。首先，描述并比较几种目前文献中提到的在 TDMA 上行信道中传输语音和突发数据的方案。我们将看到，这些接入方案都是基于时隙 Aloha 提出的。然后，将讨论 CDMA 环境下的接入控制问题。

最初，多址接入控制问题的解决方法是在 Goodman 及其同事的一系列论文中得到的（Goodman 等，1989[①]；Goodman 和 Wei，1991[②]；Nanda 等，1991[③]）。基本思想是分组存储多址（PRMA）需要一个时隙 Aloha 类型的接入并且在连续帧中传语音包。语音呼叫被预留因为语音包会周期性地不断到达，但是并不为数据分组预留时隙。数据分组必须单独地竞争时隙。先描述语音呼叫的接入和预留机制，然后再满足数据分组的接入。为进行 PRMA 性能方案所

---

① Goodman, D. J. *et al*. 1989. "Packet reservation multiple access for local wireless communication," *IEEE Transactions on Communications*, 37, 8(August), 885 – 890.

② Goodman, D. J. and S. X. Wei. 1991. "Efficiency of packet reservation access," *IEEE Transactions on Vehicular Technology*, 40, 1(February), 170 – 176.

③ Nanda, S. *et al*. 1991. "Performance of PRMA: a packet voice protocol for cellular systems," *IEEE Transactions on Vehicular Technology*, 40, 3(August), 584 – 598.

选择的帧结构满足如下假设：当接入授权后，一个用户的语音呼叫每帧需要一个时隙。

关于 PRMA 的论文，仿真时使用了如下的语音结构：通常情况下（参见第 8 章），抽样间隔 125 μs，传输速率 32 Kbps，即每个抽样点 4 比特，每 128 个抽样点组成一个 512 比特的包，每 16 ms 一个包，也就是包的长度。现在假定传输速率为 720 Kbps。第 10 章中的第三代移动通信允许这样高速率的传输，尽管实际系统中，这个速率只可能在室内环境下达到。假设每帧包含 20 个时隙，在 720 Kbps 的传输速率下，每帧有 576 比特。这意味着在每 512 比特的数据包前可以有 64 比特的控制信息和包头信息。

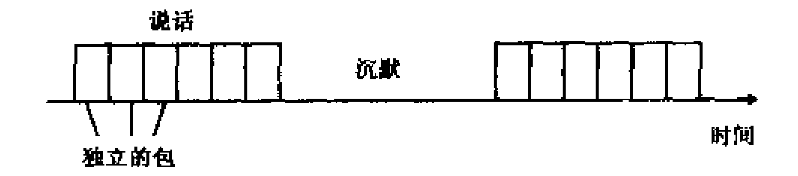
现在考虑语音接入。如前所述，PRMA 采用时隙 Aloha 接入机制，只在其上稍加改动。终端发送一个完整的包作为接入请求，若成功，基站回复应答信号（ack）。该语音终端使用的时隙就会专为该终端预留下来，基站会做出标记。这样 PRMA 中的时隙有以下两种：为正在进行的语音呼叫预留时隙和可供随机接入的时隙。从发送到接收再到应答信号的往返时间远远小于一个整帧的时间长度，甚至小于一个时隙时间。一个时隙为 0.8 ms，而以电磁波的速度，覆盖距离为 30 km 的宏蜂窝小区，信号往返时间只需 0.1 ms。即使加上基站的处理延时，还是有足够的时间在下一帧之前接收到应答信号。

在基本的 PRMA 过程之后，如果没有收到应答信号，语音终端以概率 p_v 重传下一帧，p_v 是设计参数。这种 PRMA 重传机制与一般时隙 Aloha 的区别在于，重传在一个随机间隔后进行。这种基于概率的重传方法称为 p-持续-Aloha（在下一章无线局域网的介绍中仍可见到此种重传机制）。PRMA 中的时隙预留也是预留给 Aloha 接入机制的一个例子，预留 Aloha 早已在 Aloha 相关文献中被不断研究。在 PRMA 中，语音终端最多只有两次重传机会。若还有问题，则包就被丢掉。这是因为作为实时业务，语音包必须在 100 ms 以内送到目的端。否则，谈话的对方听到的话音就难以接受，而两帧的延迟就是 32 ms。丢包势必会影响语音质量。在有线网中，关于语音分组的研究证明丢包率在 1% ~ 5% 之间还是可以忍受的。丢包率的要求通常包含在 QoS 特性中（Schwartz，1996①）。

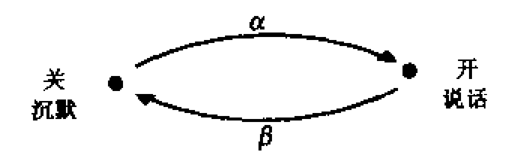
人们已经按照如上所述 PRMA 系统的具体实例进行了仿真，并将结果详细记录在以上文献中。我们仅对结果做简要讨论。仿真都在语音呼叫模型的封包过程中使用了语音激活模型。该模型很好地体现了语音的特点（Schwartz，1996），这在第 6 章关于 CDMA 系统的讨论中曾提到过。在此模型中，对讲话者的典型特征的研究表明，话音是随机的，是在说话状态和安静状态之间切换的。图 11.4(a) 表明了数据包在一段语音中周期性地产生，这一段段的语音被沉默间隔开来。图 11.4(b) 展示了两个状态的马尔科夫（无记忆）模型，该模型描述了这种语音包的产生过程。根据该模型，在任意一个很小的时间间隔 Δt 内，处于说话状态的讲话者以概率 $\beta\Delta t \ll 1$ 离开该状态，进入沉默状态。相对应地，在 Δt 间隔内，以概率 $\alpha\Delta t \ll 1$ 进入说话状态（这是连续时间模型。与其等价的离散时间模型更加适合于这个问题，因为假定在说话状态，语音与音包是周期性产生的，帧结构也是周期性的。离散模型和连续模型本质上是等价的，对于这里的计算而言，连续模型更加简洁）。稳定状态下处于说话状态的概率为 $\alpha/(\alpha+\beta)$，处于沉默状态的概率为 $\beta/(\alpha+\beta)$（认真的读者会注意到，这里两个状态的马尔科夫模型是在第 9 章讨论的更一般的多状态马尔科夫链的一个特例）。显然，平均一次讲话的长度为 $1/\beta$，平均一次沉默为 $1/\alpha$。经估测，语音突发的长度为 0.4 ~ 1.2 s，沉默期长度为 0.6 ~ 1.8 s。处于说话状态

① Schwartz, M. 1996. *Broadband Integrated Networks*, Englewood Cliffs, NJ, Prentice-Hall.

的概率为0.4。讲话者更多的时候是在沉默，这个事实早已被用在电话网中关于用户共享时隙的设计上。在一个很大的系统中，平均每2.5个用户可以共享1个时隙。由此产生了数字电话网中数字语音插空技术(DSI)，它可使2.5个用户共用一个切换。在第6章CDMA系统中，用户增益换得干扰的减小，激活的语音用户平均有60%的时间是沉默的。PRMA也希望通过语音沉默间隔来改进性能。



(a) 在一个呼叫中语音包的生成



(b) 两个状态的马尔可夫模型

图 11.4 语音中的话音激活模式

回到所引文献中的PRMA仿真，假设被准予接入某帧中的某个时隙的一个语音用户在进入沉默期后会释放这个时隙。它必须在重新开始说话时能够顺利地接入，时隙预留只会持续一个语音突发的长度。在PRMA仿真中，平均语音突发的长度是0.36 s，平均沉默间隔为0.64 s。这样，处于说话状态的概率为0.36，期望的DSI增益为1/0.36 = 2.8。一个20时隙的帧可以有56个语音通道。这是相当乐观的。仿真还显示在使用p-持续-Aloha时，p_v的最佳选择是0.3，可容纳37个语音终端，保证1%的丢包率。尽管可容纳的用户数量小于2.8倍的时隙数，但是可分配的用户数量仍然是时隙数的两倍，并且能保证他们持续整个通话时间。为什么这个值是最好的p_v呢？因为如果p_v太小，那么终端等待接入的时间太长，会导致更多的丢包；而如果p_v太大，则会导致过多的冲突，从而使系统不够稳定。

以上总结的仿真研究主要集中在单纯的语音业务上。许多后续研究考虑了既有语音终端又有数据终端的PRMA在同一上行信道竞争接入中的效应。在介绍PRMA的又一变体之前，我们只简要地介绍了其中一个仿真结果。在此研究中，Grillo和同事们考虑了包括多小区、多径、慢衰落在内的现实的操作环境(Grillo等，1993)[①]。小区是半径为1 km的六边形。路径损耗以幂指数3.5衰减，方差为4 dB的对数慢衰落。瑞利快衰落没有被考虑在内。假设基站测得的最小上行信干比可以满足基站处数据包的正确接收。有两个SIR的最小值可以考虑：9 dB和12 dB。这里的模型含捕获效应：有最高信干比的用户可以成功接入，当然该SIR值要大于最小边界值。

PRMA帧结构与上面所讲的相同，每帧16 ms，有20个时隙，以720 Kbps的比特率传输。

① Grillo, D. *et al*. 1993. "A performance analysis of PRMA considering speech, data, co-channel interference and ARQ error recovery," *Mobile and Personal Communications*, IEEE Conference Publication No. 387, December 13 - 15, 161 - 171.

每帧产生一个语音包，该包含有 512 个信息比特(话音比特)和 64 个附加比特。有一点不同在于说话-沉默的参数。平均说话长度为 1 s；平均沉默长度为 1.35 s。这样，处于说话状态的概率为 0.43；计划 DSI 增益为 2.33。一旦接入被授权，话音终端就会每帧保留一个时隙，持续一个说话长度。相反，数据终端在每次需要发送数据分组时才会请求接入，在帧与帧之间没有为数据分组预留时隙。数据包的长度为一个时隙。数据终端以相对较低的速率发送，大约为 4.8 Kbps，占空比为 50%：0.588 s 为活动期，剩余 0.588 s 不传数据。每终端的平均比特率为 2.4 Kbps。与语音包有严格的延时限制不同，数据分组如前些章所述，没有 QoS 要求。所以，在仿真中数据分组所采用的接入机制，有更长的重传时间和次数(数据包和语音包要竞争同样的时隙。一旦一个语音源获得了一个时隙的接入权，这个时隙会一直为之保留到这个语音突发结束，因此可供其他语音和数据包接入的时隙数就会减少)。另外，还有一种可变窗口的重传机制应用于数据分组。在仿真中(Grillo 等，1993)假定 30 个数据终端一直处于激活状态；通过增加语音用户的数目来增加系统流量。仿真结果显示，在 SIR 为 9 dB 时，系统可以支持 29 个语音用户，在 SIR 为 12 dB 时，系统可以支持 25 个语音用户，两种情况的丢包率都为 1%。按照上面的假设，每个成功传送的数据包平均需要传送 1.4～1.8 次。数据包额外的重传次数减小了系统的有效容量。以最小信干比等于 0.9 为例，若每个包传输 1.4 次的话，30 个用户的平均比特率为 $30\times1.4\times2.4=101$ Kbps。若屏蔽数据终端，只传送语音，则仿真显示，在理想情况下，丢包率为 1%时可容纳 37 个语音终端。这与前面的仿真结果相符合。若只传送语音，最小接收信干比为 9 dB，则只能支持 35 个移动台。若最小接收信干比为 12 dB，则只能支持 32 个移动台。以上数据显示了实际传输环境下的各种不同的折中。

在基本 PRMA 接入机制下，还可以有若干改进。其中之一是由欧洲无线研究团体的成员提出的，适用于 3G 系统，称为 PRMA++ (DeVille，1993)[①]。这种接入机制与 PRMA 的不同在于在上行链路上每一帧有两种不同的时隙。专门使用预留时隙(R 时隙)进行时隙 Aloha 类型的接入控制，在接入成功后，使用信息时隙(I 时隙)发送信息包。图 11.5(a)为一种每帧含有 36 时隙的帧结构。其中 3 个 R 时隙是确定的，剩下的 33 个是信息时隙。下行链路同样包含两种时隙：应答时隙(A 时隙)，用来确认一个成功的接入；I 时隙，用来发送含有信息的下行方向的包。这两种时隙如图 11.5(b)所示。在此接入协议中，下行链路还定义了附加的专门时隙来发送快速配置信息，在上行链路中同样定义了配置应答时隙。在此简要讨论中，我们将不考虑此附加类型。预留时隙和信息时隙的引入首先是由 Mitrou 等(1990)[②] 提出的。在此协议中，移动台通过在 R 时隙发送接入连接突发来寻求接入。该突发包含一个截短的移动台地址和发送有用信息所需的 I 时隙数目。下行链路中的 A 时隙发送成功接入的应答信息，并且确认上行链路的 I 时隙可以用来发送上行信息。只要传输是有效的，就会分配给移动台上行链路的 I 时隙。这种机制可以给一个移动台分配多个时隙，从而改变传输容量。通过在上行链路中使用有限数量的 R 时隙，可以减少 A 时隙的数量。这使得专用于解决接入问题的下行链路的容量大幅度较小(DeVille，1993)。

---

① DeVille, J. M. 1993. "A reservation-based multiple access scheme for a future universal telecommunications system," *Mobile and Personal Communications*, IEEE Conference Publication No. 387, 13 - 15 December, 210 - 215.

② Mitrou, N. M. *et al*. 1990. "A reservation multiple access protocol for microcellular mobile communication systems," *IEEE Transactions on Vehicular Technology*, 39, 4(November), 340 - 351

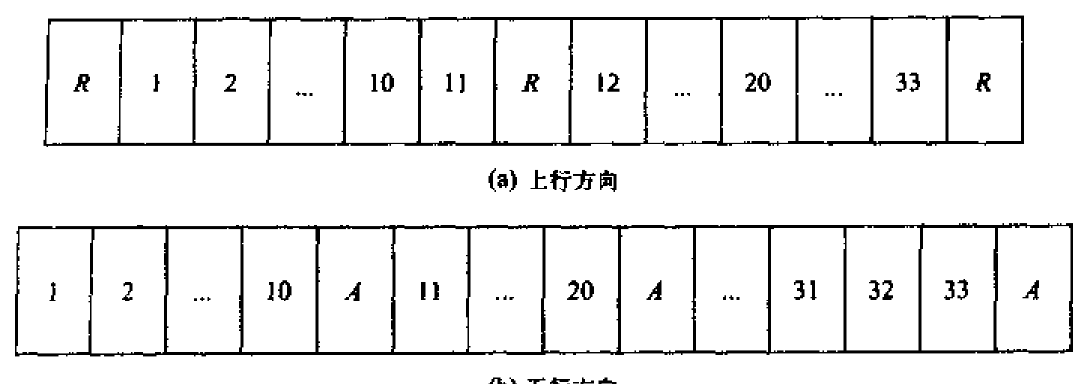


图 11.5 PRMA++帧结构示例

有人认为，如果每帧的时隙数确定，则可以为最小化接入延时找到一个最优的每帧中 R 时隙的数目。接入成功后，接入延时由两部分组成：竞争延时和等待信息时隙变为空闲的延时。若 R 时隙较少，则竞争延时会较长。随着 R 时隙数目的增加，竞争延时会缩短，但 I 时隙数也会随之减少。而 I 时隙数的减少会导致接收到 I 时隙的延时增大，同样会使接入延时增大。这样，最优的 R 时隙数是与系统中移动台数目有关的。比如，DeVille(1993)的仿真显示，在语音分组移动通信系统中，假定每帧 5 ms，含 72 个时隙，128 个移动台，16 个 R 时隙，那么最小接入延时为 20 ms(4 帧)。其中，设定平均语音突发长度为 1.41 s，平均沉默长度为 1.78 s。这样，处于说话状态的概率为 0.44，期望 DSI 增益为 2.27。在电路交换模式下，实际增益为 128/72 = 1.78，可以把它与上述 PRMA 仿真数据做个粗略的对比，但它的帧结构有所不同。这样，在下行链路中只需 16 个 A 时隙，其余 56 个时隙都可以传信息(这里再次忽略了其他用途的时隙，如配置时隙)。如果接入延时可以是 30 ms(接近于 PRMA 仿真中的 32 ms 最大值)，可以容纳 136 个移动台，每帧中 R 时隙的数目降到了 12 个(DeVille，1993)。

在文献中还提到了 PRMA 和 PRMA++的很多其他扩展。例如，称为无线业务整合(SIR)的接入协议，就是在 PRMA++的基础上将数据业务和语音业务整合在一起(Anasti 等，1996①，1997②)。在此方案中，上行链路 R 时隙被分为两种：VR 和 DR。前者与 PRMA++中的一样，用于竞争接入，还可用来为语音业务申请带宽。后者用来向基站提出无竞争的数据带宽请求：当数据终端被成功接入后，基站会周期性地轮询这些移动台来决定为其分配多少带宽。基站以轮询、无竞争的方式将 I 时隙分配给各个数据终端，专门用来传送数据的 I 时隙就这样提供给数据用户。另外，每帧中空闲的语音 I 时隙也会分配给此帧中的数据使用。这种允许数据使用语音时隙的机制称为“可移动边界”，早已被用在了有线的链路分配机制中。这样，在接入成功后，SIR 接入机制可以以无竞争的、轮询的、不影响语音信息时隙的方式给数据用户分配信息时隙。

另一种改进的综合接入协议为帧预留多址接入(FRMA)，它对于基本 PRMA 的改进有以下两个方面：一方面，它在保证语音业务的丢包率始终为 1%的情况下尽量增多每帧中分配给数

① Anasti, G. *et al*. 1996. "A bandwidth reservation protocol for speech/data integration in TDMA-based advanced mobile systems," *Proc. IEEE Infocom' 96*, San Francisco, CA, March, 722 – 729.

② Anasti, G. *et al*. 1997. "A contention/reservation access protocol for speech and data integration in TDMA-based advanced mobile systems," *Mobile Networks and Applications*, 2, 3 – 18.

据的时隙数；另一方面，它将基站广播的应答信息减少为每帧一个，以此来减少移动台的接收处理，延长电池寿命（Narasimhan 和 Yates，1996[①]）。特别地，在上行链路中，时隙被分为 3 种：为语音信息包保留的 R 时隙，一个激活的语音移动台每帧只分得一个 R 时隙；为语音用户竞争设置的 V 时隙，以及为数据竞争和数据传输设置的数据时隙（此处的命名与 PRMA++ 有所不同：这里的 V 时隙相当于 PRMA++ 中的 R 时隙；这里的 R 时隙相当于 PRMA++ 中的 I 时隙）。这里也采用了可移动边界策略，将未被占用的语音时隙分配给数据业务。如果每帧共有 N 时隙，数据时隙的数目即为 $N-R-V$。在这里，基站自适应地计算所需的 V 和语音 p 持续概率 p_v，为了最大化每帧中的数据带宽，并将其发送到所有的语音终端。前面已提到，最大化数据带宽必须在保证语音业务丢包率在 1%以内的前提下进行。为此，基站需要在每帧开始时知道正在通话的语音移动台数目 M 和处于语音突发状态的数目 R，每个用户使用一个语音信息包来传送（$R \leq M$）。V 和 p_v 都是 R 和 M 的函数。

使用这种接入方式的效果怎样？仿真显示，由于采用了优化方法，它的数据终端性能远远好于纯 PRMA（Narasimhan 和 Yates，1996）。考虑与上述系统参数完全相同的系统：每帧 16 ms，含 20 个时隙；数据率为 720 Kbps，每个语音呼叫占用带宽 32 Kbps。每时隙有 64 比特包头加在 512 比特语音信息中；平均说话突发 1 s，平均沉默长度 1.35 s（说话突发和沉默长度与前所述略有不同）。在 PRMA 的仿真中没有考虑衰落效应，结果是随着语音用户数的增加，能为数据终端提供的带宽（时隙）的比例线性减小。当系统中有 32 个语音终端时，平均只有 20%的带宽（即 144 Kbps）可供传输数据。在语音业务的 PRMA 系统中，一个基站最多接入 37 个用户。在既传送语音又传送数据的整合 PRMA 系统（Grillo 等，1993）中，若有 30 个语音用户，则有 72 Kbps、10%的带宽提供给数据终端，相当于 30 个速率为 2.4 Kbps 的终端。考虑另一个例子，系统中有 20 个语音终端和 20 个数据终端。在 FRMA 机制下，50%的带宽（360 Kbps）可传送数据，每个速率为 18 Kbps。而且，在 FRMA 机制下，数据包在无线信道传输队列中的平均延迟大大小于 PRMA 机制。考虑在 20 个语音终端和 20 个数据终端情况下，每个终端的数据率提高时的数据包平均延时。当每终端数据率为 1000 bps 时，PRMA 机制下的延时为 400 ms，而相应的 FRMA 延时则小于 100 ms。在数据率进一步增加时，两种机制的队列延时都会沿着典型的队列曲线有所增加。PRMA 机制在每终端 1400 bps 时达到饱和，此时延时为 800 ms。而 FRMA 机制可以达到 18 Kbps 的数据率，同时延时小于 800 ms（Narasimhan 和 Yates，1996）。FRMA 接入机制能提供如此显著的性能优势的代价在于，语音业务的丢包率即使在系统负荷很轻时仍然为 1%。这一数据已经写进了优化程序。而在 PRMA 中，1%的丢包率只是一个极限值，即系统的实际丢包率小于等于 1%。在负荷较轻时，PRMA 系统的丢包率会远远小于 1%。

我们将要讨论的最后一种基于 TDMA 的接入机制是分布式队列请求更新多址（DQRUMA），它是为固定速率的无线局域网而设计的一种接入协议（Karol 等，1995a）[②]。在另一篇论文中，作者还研究了在蜂窝小区较大时 DQRUMA 的改进方法（Karol 等，1995b）[③]。该协议试图更有

① Narasimhan, P. and R. P. Yates. 1996. "A new protocol for the integration of voice and data over PRMA," *IEEE Journal on Selected Areas in Communications*, 14, 4 (May), 621 – 631.

② Karol, M. J. *et al*. 1995a. "An efficient demand-assignment multiple-access protocol for wireless packet (ATM) networks," *Wireless Networks*, 1, 5 (October), 267 – 279.

③ Karol, M. J. *et al*. 1995b. "Distributed-queueing request update multiple access (DQRUMA) for wireless packet (ATM) networks," IEEE International Conference on Communications, ICC'95, Seattle, WA, June, 1224 – 1231.

效地利用系统容量。上行链路的帧除了含有通常包长的信息帧外,还有一种微时隙(mini-slot)接入请求(RA)信道。这种微时隙 RA 信道会比正常时隙的竞争信道占用更少的资源。移动台从空闲信道中随机选择一个 RA 信道,发送一个 b 比特长的移动接入标识符来竞争接入。b 比特数取决于一个基站可支持的移动台数。帧与帧之间 RA 信道数依据移动台接入请求的数目动态变化。当流量较小时,基站可以给每帧分配更多的 RA 信道,结果通过下行消息通知给各移动台。

在下行链路的帧结构中,为每个上行 RA 信道分配一个应答信道。若一个 RA 信道的接入成功,则移动台会在下行的 ACK 信道中识别出自身的标识符,并在上行方向的下一时隙发送有用信息。这种机制主要应用在无线局域网中,因为无线局域网中的往返时间和传输延时都比较小。如上所述,还有工作表明 DQRUMA 协议在蜂窝小区较大的系统中是如何工作的(Karol 等,1995b)。如果移动台还有另外的信息包需要发送(它的队列不是空的),则它可以在传送信息包时将一个 1 比特的捎带请求标识置位,以此来通知基站。接下来的接入就会是无冲突的(后面会简要介绍一种在 CDMA 系统下的 DQRUMA,它采用了 p 比特的捎带请求位,用来指示移动台接下来还要发送的信息包的数目。在 RA 微时隙的初始接入请求中也会保留这几位)。它的竞争接入机制采用了一种改进的时隙 Aloha 接入机制,用户的接入概率是可变的(Karol 等,1995b)。仿真显示,这种动态时隙 Aloha 和所提出的另一种算法效果相当好,可以得到接近理想接入机制的包延时和吞吐量性能。这种理想接入机制假设基站准确地知道每个时隙所有的接入请求信息。捎带请求比特机制减少了随机接入的请求数,因为如果一个移动台的缓存是非空的,它就没有必有使用第 1 次接入后的随机接入。仿真显示,如果突发的长度由 1 个包长增加为两个包长,则包延时-吞吐量性能就会明显提高(这种机制类似于 PRMA 中的语音接入,不同的是,在 PRMA 的语音接入中,基站需要一帧的时间来发现以前分配的语音信息时隙变空了,从面知道该语音移动台不再处于说话状态;而这里的移动台是否还有包等待传输是从当前帧的传输中知道的)。

CDMA 接入机制

在这一节中,我们将详细介绍由 PRMA 机制演化而来的应用于 CDMA 系统的接入策略。这些策略通过相互重叠的时隙使用不同的码字来传输,应用在第三代 CDMA 系统中的方法将在第 10 章讨论。这些方法是直接应用于这些系统的。

从 CDMA 版本的 DQRUMA 入手,曾在上一节讨论 DQRUMA 时提到(Liu 等,1996)[①]。这里的多址方式是多码 CDMA,是第 10 章讲到的通过使用正交扩频码得到更高传输速率的可变速率 CDMA 的一种形式。正交扩频码通过"子码连接"过程得到。在此过程中,每个移动台分配一个基本伪随机码,这些码不是完全正交的,通过用一系列的 Walsh 码对基本码进行操作得到正交扩频码(Chin-Lin, I. 和 Gitlin, 1995c[②]; Liu 等, 1996)。

在基于 TDMA 的 DQRUMA 中,使用了时隙 Aloha 为各个帧分配 RA 微时隙。移动台使用惟一的正交码或 b 比特短识别码来请求接入。当两个以上的移动台选择相同的正交接入码时

① Liu, Z. *et al*. 1996. "A demand-assignment access control for multi-code DS-CDMA wireless packet (ATM) networks," *Proc. IEEE Infocom' 96*, San Francisco, CA, March, 713 – 721.

② Chin-Lin, I. and R. D. Gitlin. 1995. "Multi-code CDMA wireless personal communication networks," IEEE International Conference on Communications, ICC95, Seattle, WA, June, 1060 – 1064.

就会发生冲突。接入中含有一个 p 比特数代表该移动台欲发送的包的个数。基站为每一个请求接入的接收端配置了一个指定的码字,可使多个移动台的接入请求在一个下行微时隙信道中得到应答。使用不同码字的多个终端的接入请求就可以得到识别和应答。这种同时允许多个接入请求的方式就是 CDMA 系统与 TDMA 系统相比的最显著特性。但是回忆第 6 章内容可知,CDMA 接入技术的带宽远远大于 TDMA(上节所述的各种 TDMA 接入方式的性能结果都是针对一个频点信道而言的)。另外,同时接入多个用户也导致了干扰的增加。

在多码 CDMA 版本的 DQRUMA 中,移动台直到收到了基站发出的允许发送指示时才能发送信息包。该指示作为多发允许指令中的一种,与基于 TDMA 的 DQRUMA 中的应答信号极为相似,只是在发送允许微时隙信道时使用了一个多码 CDMA 过程。该消息会告知移动台码的个数、传输速率、在下一帧发送包时所需的基本伪随机码和所需使用的发射功率等级。鉴于码字是每个时隙分配一次,并且可以复用,这就大大减少了所需码字的总数(Liu 等,1996)。

这种多码 CDMA 版本的 DQRUMA 流量的仿真结果与上一节总结的基于 TDMA 的 DQRUMA 基本相同(Liu 等,1996)。对于携带多个包的数据突发,通过增加每时隙中的突发个数而得到的延时-吞吐量曲线与理想接入曲线极为接近。之所以接近理想情况,是由于在 DQRUMA 中使用了捎带请求接入机制。实际上,关于一个包的突发仿真显示,DQRUMA 的性能只有在业务流量很大时才接近理想曲线,因为流量大时,移动台的缓冲区大多数时间不是空的,这样才能使捎带请求接入机制派上用场。

下面来介绍一种更接近于 PRMA 的适用于语音和数据的 CDMA 接入机制(Brand 和 Aghvami,1996)[①]。正如在早期的基于 TDMA 的 PRMA 中,欲发送信息的终端以概率 p_v 发送语音,以概率 p_d 发送数据。在帧与帧、时隙与时隙之间,这个允许概率随时变化。每个时隙,基站计算一次,并通过下行链路发送到移动台,便于下一帧使用。在论文中写道,p_d 始终是 p_v 的一小部分,通常是 p_v 的 0.1 或 0.2,这样只有 p_v 需要计算。选择 p_v 的目标在于在语音丢包率限制在特定数目的情况下,通常是 0.01 或 0.02,极大化吞吐量。这个目标等价于控制同一时间某些帧内每帧周期性用户的数目。周期性用户是指在一段连续的帧中每帧传送一个信息包的用户。语音包的发送有这样一个例子:这里的数据包可以是周期的也可以是随机的,但每次只能传送一个包。在 PRMA 中,一个周期性用户如果成功接入某个时隙,这个时隙就会一直为它保留,直到包发送完毕。在给定时隙中,一个最优的预留数目的选取与能够满足最大吞吐量的数目是相关的。显然,如果出于预留状态的用户数很大,允许率就会下降。事实上,如果一帧中最优的用户数到达了,允许概率就会是 0。当用户数降到低于最优值时,允许概率才会增加。为了计算方便,允许概率和用户数量之间的函数关系被定义为两个线性部分,开始的曲线下降得比较平缓,当用户为 0 时达到最高的允许概率,但随着用户数的增加,允许概率逐渐下降,到达某个拐点处后,随用户的增加迅速降到 0。

对这种机制在单蜂窝小区和六边形多蜂窝小区两种情况下的性能进行了仿真(Brand 和 Aghvami,1996)。与前面的蜂窝系统相同,语音源速率为 8 Kbps。假设平均说话突发为 1 s,平均沉默长度为 1.35 s,与 Grillo 等(1993)所述相同。相应的 DSI 增益为 2.35,帧长为 20 ms,每帧 20 时隙。每个语音源每帧有 160 个信息比特。这些比特加上包头,经过编码,构成将在一个

① Brand, A. E. and A. H. Aghvami. 1996. "Performance of a joint CDMA/PRMA protocol for mixed voice/data transmission for third generation mobile communications," *IEEE Journal on Selected Areas in Communications*, 14, 9(December), 1698 - 1707

给定时隙中发送的语音包。如果包不能在 20 ms 内发送就会被丢弃。首先考虑单小区仿真结果。在只传话音业务时，在最大丢包率为 1%～2%时，最佳的每时隙含有周期源数目为 9。最佳的允许概率曲线在没有用户时起始端为 0.3，然后线性减少到每帧有 6 个用户时的 0.25，随后迅速减少到每帧有 9 个用户时的 0。在较宽松的最大丢包率为 2%的限制下，同时可支持的呼叫数为 379。如果这个限制减少到 1%，则可支持的呼叫数下降到 358。这些数可以与使用 DSI 计算得到的 9×20×2.35=435 相对比进行观察。如果使用纯随机接入，设 $p_v=1$，在丢包率为 2%的限制下同时通话的用户数为 227。这样，控制接入概率的技术可以获得 67%的性能增益。在丢包率限制为 1%时，性能增益为 78%。干扰模式绘制出了每时隙的用户数随时间变化的曲线，发现在随机接入模式下，该数目以 8 为均值大范围地波动，而在控制接入下，该数目只在 7 到 9 之间变化。

在业务中加入单个包的数据业务的影响已经通过仿真被估测出来，在仿真中通过增加数据终端的数目来计算在丢包率为 2%的限制下可支持的语音终端的数目（Brand 和 Aghvami，1996）。假定数据终端的比特率为 3400 bps，与语音终端平均比特率相同(8 Kbps×0.43)。当数据终端增加时，可容纳的语音终端如期减少。由于突发性的增加，网络吞吐量随之减小。但是与随机接入（Brand 和 Aghvami，1996）相比，在容量上还是有很大的提高。

在六边形小区结构中的仿真结果与接入控制的结果相似。与单小区结构相比，在给定丢包率限制下能同时接入的语音数有所减少：小区间干扰大大增加了系统的总干扰（回忆在第 6 章中 CDMA 容量的计算，显示出小区间干扰远大于小区内干扰）。但是，可控接入还是远远优于纯随机接入（Brand 和 Aghvami，1996），改进的数据取决于所选择的扩散模型。以前面章节所述的通用$1/d^4$ 模型为例，在丢包率为 1%的限制下，使用可控接入时单小区中能同时支持 358 个用户，现在降低到 280 个，相应的随机接入数为 160 个，可控接入性能提高了 75%，这与前面所述的单小区情况类似。

11.3 基于分组的蜂窝系统的调度

如本章引言所述，为蜂窝系统中包传输所设计的资源分配，如第 10 章所述的第三代蜂窝系统和将来的第四代系统，需要对上行和下行方向上的数据包进行合理的调度。调度有两层目的：一是为了每一个用户在信道条件允许的情况下提供适当的 QoS 保证，诸如最大包延时、相关的吞吐量保证、丢包率和其他 QoS 性能保证；二是使带宽、容量等资源得到最大利用。在第 10 章的 10.2 节，cdma2000 的结尾部分简要介绍了资源分配的必要性，讨论了测得的 SIR 和在此 SIR 下可能的数据率的关系。为了给出 QoS 的要求是如何影响资源分配的例子，我们引用了一篇文献，说明了数据延迟时间的改变使系统总体吞吐量有所提高。

图 11.6 用图示来表明调度的含义。这是个十分抽象的模型，有 N 个数据缓冲区，用 $q_i(t)$表示在 t 时刻第 i 个队列包含的包的数量，用 $r_i(t)$表示第 i 个队列发送或释放包的速率。数据缓冲区可以显示上行方向上等待发送的数据包。在相反的方向上，基站到 N 个移动台之间的数据包由基站控制；它们同样可以表示基站里与每个移动台相关联的缓冲区中存在的下行方向上需要发送到每个移动台的包的数目。调度问题，包括上行和下行，就是要决定在 t 时刻要发送哪个或哪些数据包。图 11.6 所示的包传输速率 $r_i(t)$就是运用调度算法的结果。它们将作为解的一部分写入调度算法。在下面的第 1 个例子中，我们将看到通过使用算法，该

速率得到了有效的提高。在此例中，对于给定的移动台，期望数据率是由测量得到的 SIR 决定的。这就是第 10 章所采用的方法，在应用了调度算法后，实际的数据率可能有所不同。

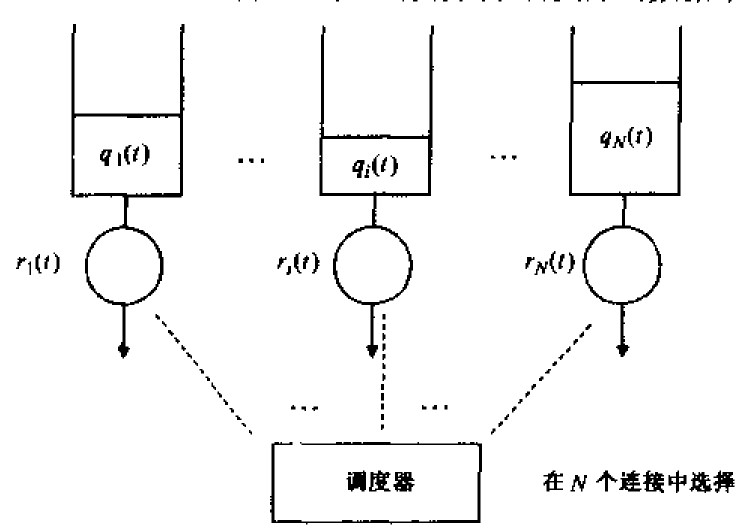


图 11.6 多个数据流的调度

调度算法通常是针对一些期望的性能目标来设计的，这些目标可以是在最不利的传播环境下满足每种数据类型的 QoS 要求，像刚才所述或 10.2 节中的例子所示。目标还可以是在给定的无线链路上寻求最大的系统吞吐量，上面也提到了。调度算法的选用显然还要考虑给定业务流的类型和特性(其中的一些特性可能已经在 QoS 要求中体现出来了)。也就是说，对于突发数据流的调度需要考虑业务的突发特性，而规则的连续类型的业务调度需要反应出其连续特性。宽带业务，比如网站中下载到移动台上的多媒体业务，显然与数据率小得多的电子邮件业务的处理方法完全不同。

在系统设计中，调度任务在不同时隙间、不同帧间都有所不同，图 11.6 中示出了在 t 时刻做出了哪种选择。对于一个给定的移动台，在上行或下行方向上，数据包的实际传输速率，在 CDMA 系统中由分配的码字所确定，在 TDMA 系统中由分配的时隙数所确定。第 10 章给出了在这两种情况下的例子。在一个给定的系统中，调度过程可以由基站单独完成，也可以由位于控制若干个小区的网络控制器中的调度器完成(WCDM，2000)[①]。无论哪种方式，要实行调度算法都需要使用第 10 章所述的链路控制信道来发送必要的信息，比如各移动台测得的功率和 SIR、移动台中等待发送的包的特性和数量等(在上节所述的 DQRUMA 系统中，上行链路还要包括移动台队列的长度)。

如图 11.6 所示，通过在多个业务流中进行调度来满足一定的性能目标的概念已经在有线分组网络中进行了很久的研究，可参阅 Zhang(1995)[②]的著作中分组网络调度机制的综述。为无线系统设计提出的一些调度算法都是在原始的有线系统上做出的变形。在 Fattah 和 Leung(2002)的著作[③]中有关于 TDMA 和 CDMA 无线网络调度机制的综述。该综述还包括这些算法首次在

① WCDM.2000. *WCDMA for UMTS*, ed. Harri Holma and Antti Toskala, John Wiley & Sons.

② Zhang, H. 1995. "Service disciplines for guaranteed performance service in packet-switching networks," *Proc. IEEE*, 53(October), 1374 – 1396.

③ Fattah, H. and C. Leung. 2002. "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wireless Communications*, 9, 5(October), 76 – 83.

起初的有线网络中应用的文献参考。这一节，为简单起见只选择了其中的一些调度算法并且着重关注它们的调度原理。从第 10 章介绍的 cdma2000 1X 高数据率(HDR)系统所采用的调度算法入手。这种算法称为比例公平(PF)调度算法(Rental 等，2002)[①]。该算法基于时隙，操作简单：选择用户队列中所需传输速率与队列平均传输速率比值最大的包进行传输。这种关系一旦确定，就打破了随机性。具体地，令 $DRC^i(n+L)$ 代表第 i 个用户在时隙 $(n+L)$ 基于对 SIR 的测量而得到的所需数据率。如 10.2 节中关于 cdma2000 所述，用户速率的选择是基于所测得的 SIR 值的，它们是在保证误包率为 1% 时的最大速率(Bender 等，2000)[②]。令 $R^i_{av}(n+L-1)$ 代表用户 i 经过测量以前的时隙而得到的该用户所接收到的平均传输速率。在第 $(n+L)$ 个时隙，被传送的用户的 $DRC(n+L)/R_{av}(n+L-1)$ 值最大。比例公平算法在时隙 $(n+L)$ 时选择用户 j 进行发送，如果 j 有如式(11.11)所示的最大比例的话：

$$\text{PF:}\quad \max_i\left[DRC^i(n+L)/R^i_{av}(n+L-1)\right]\quad 1\leqslant i\leqslant N \tag{11.11}$$

该算法欲满足有最高请求速率的用户的需求，但是同时引入了一种公平原则，这种公平原则由经平均传输速率规一化后的请求速率来衡量。这样，平均传输速率高的用户的优先级就会比平均传输速率低的用户靠后。该算法试图最大化信道吞吐量(选择有最大请求速率的用户可以清楚地证明这一点)，但同时引入了一种规一化的公平机制，来防止用户贪婪地占用信道。

这样就需要使用平均算法来计算平均速率，Rental 等(2002)提出了一种一阶递归的更新方法

$$R^i_{av}(n)=\left(1-\frac{1}{t_c}\right)R^i_{av}(n-1)+\frac{1}{t_c}r_i(n-1) \tag{11.12}$$

其中，$r_i(n-1)$ 代表用户 i 在前一时隙 $n-1$ 收到的实际发送速率，这就是图 11.6 所示的发送速率。t_c 是一阶平均过程每次平均的时隙数。在 Rental 等(2002)的仿真中这个数是 1000，由式(11.12)可以看到，平均传输速率随时间的变化将会很缓慢。

例子

在 Andrews 等(2001)的著作中，一个基于此算法的、简单的、两用户的例子显示了该算法与轮询方式的系统资源分配相比，有效地提高了用户和系统的吞吐量。在轮询方法中，系统依次为用户分配带宽，没有优先级的区别。在采用比例公平算法的这个例子中，假设用户 2 的平均信道质量比用户 1 好，由 SIR 的测量值所决定的传输速率有两个：153.6 Kbps 和 307.2 Kbps，每种速率有 50% 的机会取到。用户 1 由于信道质量较差，只能以 76.8 Kbps 和 153.6 Kbps 两种速率传输，每种各占 50%(这是 cdma2000 1X 所提供的数据率，详见 10.2 节和 10.3 节)。读者可以自行证明，在轮询方式下，信道平均吞吐量为 172.8 Kbps，用户 1 的平均吞吐量为 57.6 Kbps，用户 2 为 115.7 Kbps。

下面引入比例公平机制，式(11.11)中所需的平均速率 R_{av} 可以通过将每个用户的两种速率作平均得到，即 $R^1_{av}=115.2$ Kbps，$R^2_{av}=230.4$ Kbps。这里没有必要使用式(11.12)的递归公式，稳态的平均值已经足够。在两个用户的情况下采用比例公平算法时，可能有 4 种等概率的情况出现：有两种平局情况，$76.8/115.2=153.6/230.4=0.67$ 和 $153.6/115.2=307.2/230.4=1.33$。第 1 种情况的平均传输速率为

---

① Rental, C. H. *et al*. 2002. "Comparative forward link traffic channel performance evaluation of HDR and 1EXTREME systems," Proc. *55th IEEE* Vehicular Technology Conference, VTC Spring, 160 – 164

② Bender, P. *et al*. 2000. "CDMA/HDR: a bandwidth-efficient high-speed service for nomadic users," *IEEE Communications Magazine*, 38, 7(July), 70 – 77.

0.5(76.8 + 153.6) = 115.2 Kbps;第 2 种情况为 230.4 Kbps,是第 1 种情况的两倍;第 3 种情况是用户 2 以 307.2 Kbps 的速率胜出,如 307.2/230.4 > 76.8/115.2;第 4 种情况是用户 1 以 153.6 Kbps 的速率胜出,如 153.6/115.2 > 153.6/230.4。对这 4 种情况进行平均,得到总的信道吞吐量为 201.6 Kbps,这明显大于轮询方式的 172.8 Kbps,每个用户的平均吞吐量也有所提高,分别是 67.2 Kbps 和 134.4 Kbps。具体细节读者可自行证明。

Rental 等(2002)所做的关于 HDR 系统的大规模的仿真显示,比例公平机制在 3 扇区的小区和步行环境下当移动台数目增加时可以使系统吞吐量有显著的提高。路径损耗、阴影衰落和瑞利衰落都被包含在仿真中。在每扇区 16 个用户的情况下,每个用户的平均吞吐量为 100 Kbps,标准差为 55.6 Kbps。标准差提供了该算法的公平性量度。在包含了车载环境的另一个仿真中,当车速为 60 km/h 时,还可以观察到使用了 PF 算法的边缘增益(Rental 等,2002)。

比例公平算法是 Andrews 等(2001)[①] 提出的一种更通用的调度算法的一个特例。这种更通用的算法,称为改进的最大加权权重延时优先(M-LWDF)算法,它的特点在于保证所有用户的所有队列都处于稳定状态,没有队列丢包(这里假设队列为无限长,如果要设置一个有限值,就相当于在超过这个值后,队列饱和在此值不变,新到来的包直接丢弃)。当且仅当下式取最大值(Andrews 等,2001)时,该算法将在 t 时隙调度队列 j 中的包

$$\text{M-LWDF:} \qquad \max_i[\gamma_i W_i(t) r_i(t)] \qquad 1 \leqslant i \leqslant N \tag{11.13}$$

这里 γ_i 是一个指定的常数,它的值将在下文讨论;$W_i(t)$是位于队列 i 头部的包已经等待的时间;$r_i(t)$是前面所述队列 i 的平均传输速率。部分或全部用户中的等待时间 $W_i(t)$可以用图 11.6 所示的队列长度 $q_i(t)$代替(Andrews 等,2001)。那么,怎样选择常数 γ_i 呢? 这取决于业务的类型和 QoS 要求。Andrews 等提供了一个有延迟限制的业务的例子,比如语音或视频这样的实时业务。对于这种业务,队列 i 的延迟限制可以表示为下式:

$$\text{Prob}[W_i > T_i] \leqslant \varepsilon_i \tag{11.14}$$

这里 T_i 是延时限制;ε_i 是指定概率。对于实时语音,可以取 $T_i = 50$ ms, $\varepsilon_i = 0.01$,也就是说,有 99%的概率能够满足 50 ms 的延时限制。Andrews 等建议,在给定了延时限制为式(11.14)后,将常数 γ_i 写为 a_i/R_{av}^i。其中,R_{av}^i就是 PF 算法中已经用过的第 i 个用户的平均传输速率,a_i 被定义为$(-\log\varepsilon_i)/T_i$。比如,对于用户 i,取 $T_i = 50$ ms, $\varepsilon_i = 0.01$,得到 $a_i = 0.04$。如果 ε_i 增加到 0.1,限制条件放宽,那么 $a_i = 0.02$。用户 i 的优先级就相应地降低了。对于小区中的某个用户的传输优先级可以通过选择参数 a_i 的值来改变。利用 $\gamma_i = a_i/R_{av}^i$的关系,M-LWDF 调度算法定义为在时隙 i 内选择用户 j,如果用户 j 在 N 个移动台中下式的结果最大:

$$\text{M-LWDF:} \qquad \max_i\left[a_i W_i(t) r_i(t)/R_{av}^i\right] \qquad 1 \leqslant i \leqslant N \tag{11.15}$$

下面研究为何这种形式的 M-LWDF 调度算法与式(11.11)描述的 PF 算法是类似的。这里有 3 个参数影响了在 t 时刻哪个用户被选定:a_i 和 $W_i(t)$可看做提供了 QoS 延时限制要求,比值 $r_i(t)/R_{av}^i$表示最大化信道总吞吐量的能力,而并没有像 PF 算法那样过分地惩罚低速率用户。虽然上述讨论的 M-LWDF 算法的例子是针对实时业务的,但它在非实时业务中同样适用。在

① Andrews, M. et al. 2001. "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, 39, 2 (February), 150 – 153.

这种情况下，可以忽略 a_i（将其对所有用户取常数）。等待时间 $W_i(t)$ 或其等价序列长度 $q_i(t)$ 保证了每个队列中的包不被过分地延迟。

M-LWDF 算法是与等待时间 $W_i(t)$ 或其队列长度 $q_i(t)$ 结合在一起的。比其他的包都等得更久的包将有更高的优先级。这种对于等待时间长的包的偏好与一类称为"最早时限优先"(EDF)的调度算法有关。这种算法以及它的变形早已在有线网络中进行了很多研究。EDF 算法在每个时隙选择位于队列头部的包和离延时限最近的队列进行发送。用符号来表示，并且参考图 11.6，这意味着选择队列 i 中的位于队首的包，如果该队列 $(T_i - W_i)$ 只是所有 N 个队列中最小的。这里 T_i 是式(11.14)中定义的延时限。在文献 Varsou 和 Poor(2002)[1]以及 Solana 等(2002)[2]中，研究和比较了适用于 CDMA 系统的各种版本的改进算法。后一篇文献中，当一个包到达发送队列时，使用了时间打戳的方法来记录它们到达队列首的时间。时间打戳过程是由其他学者在研究基于包的等待时间的调度算法时提出的。

在 Varsou 和 Poor(2002)引用的一份未出版的 1999 年 9 月的贝尔实验室备忘录中显示，"最早时限优先"(EDF)类型的调度算法关于下行调度机制提供了一种很好的性能。Varsou 和 Poor 将未公开的工作扩展如下：在原始工作的每个时隙中，选择离最后期限最近的在队列头部的下行包进行发送，所有的功率都分配给了这个发送。Varsou 和 Poor 允许将剩下的功率分配给下一次要传输的包，这种调度机制称为"功率最早时限优先"(PEDF)。还有一种改进，称为"功率最早时限优先公平"(PEDFF)机制，将剩余的功率分配给所有处于激活态的用户。它在调度算法中引入了一种公平机制。这两种改进的 EDF 算法都提供了更好的性能，提高了功率利用率(Varsou 和 Poor,2002)。PEDFF 算法对于突发业务和混合业务的性能比 PEDF 算法好。通过使用一种偏好包长较短的传输，惩罚错过延时限的包的改进调度算法可以得到更高的性能(Varsou 和 Poor,2002)。

在这一节的最后将介绍一种由有线网络中的调度机制演化而来的针对 CDMA 设计的公平调度机制。首先描述有线网络中的形式。这种机制称为通用处理器共享(GPS)，它是由 Parekh 和 Gallager(1993)[3]首先提出和分析的。Schwartz(1996)的著作中有该方法的概要。GPS 算法是一个理想的算法，并不适用于实际应用，因此 Parekh 和 Gallager(1993)还讨论了一种实用的扩展方法称为分组 GPS(PGPS)，也称为权重公平队列。这里为了简化讨论，只介绍 GPS 算法。GPS 本质上是一种改进的轮询机制，它能够在多个用户间公平地分配链路容量(带宽)。在执行轮询时，它通过计算已经分配给各用户的容量来达到公平的目的。具体来考虑图 11.6 中的 N 个队列，让它们共享一个链路容量 C。在 GPS 的相关文献中，一个队列以及与其相关的传输元素称为一个连接。想像在轮询形式下的各个连接，假设第 i 个连接被分配的权重为 ϕ_i，那么就要保证它的业务速率为 g_i，由下式给出

$$g_i = \phi_i C \Big/ \sum_{j=1}^{N} \phi_j \tag{11.16}$$

---

[1] Varsou, A. C. and H. V. Poor. 2002. "Scheduling algorithms for downlink rate allocation in heterogeneous CDMA networks," *Journal of Communications and Networks*, 4, 3(September), 199 - 208.

[2] Solana, A. H. *et al*. 2002. "Performance analysis of packet scheduling strategies for multimedia traffic in WCDMA," *Proc. 55th IEEE Vehicular Technology Conference*, VTC 2002, Spring, 155 - 159.

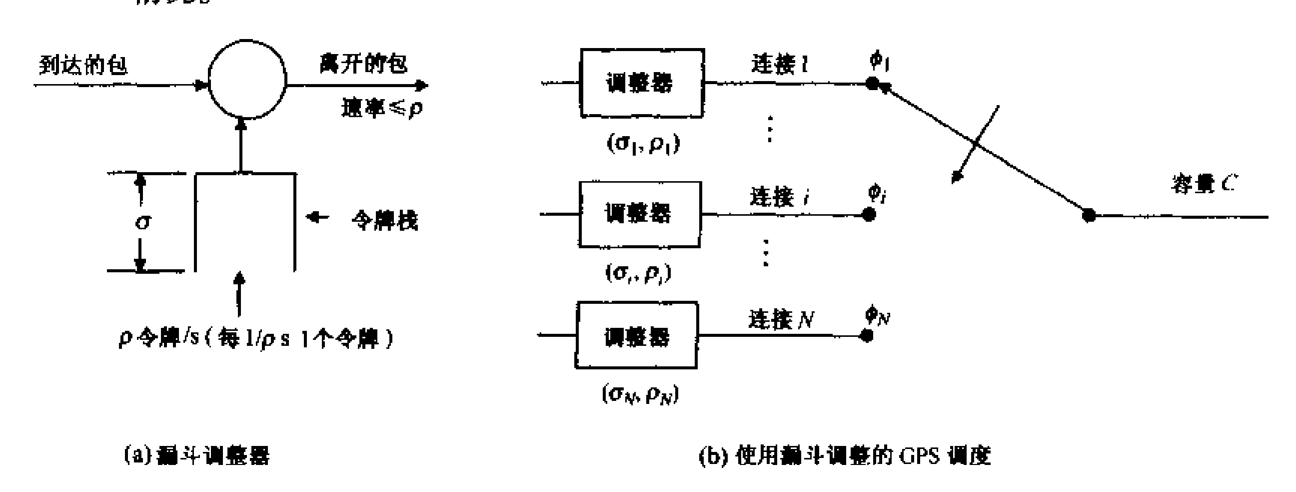
[3] Parekh, A. K. and R. G. Gallager. 1993. "A generalized processor-sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Transactions on Networking*, 1, 3(June), 344 - 357.

图 11.7(b)给出了 GPS 调度的基本思想，其中调整器(regulator)将在后面介绍。GPS 机制还定义了保存功能，这意味着每个机会都有包传输。这样，如果轮到一个连接传输数据而它暂时没有业务可传时，就会分配其他连接来传送业务。这样，在长度为 T 的时间间隔内，第 i 个连接传送的业务量(用 S_i 表示)服从如下不等式

$$S_i/S_j \geqslant \phi_i/\phi_j, \qquad j = 1 \cdots N \tag{11.17}$$

可以看到，GPS 提供了如下两个保证(Parekh 和 Gallager, 1993)[①]：

1. 令 R_{av}^i 代表第 i 个连接的平均业务速率。如果 $R_{av}^i \leqslant g_i$，那么 g_i 是可以独立于其他连接被保证的。
2. 每个连接上的最大业务延时是有界的，它只基于自身的特性，而独立于其他连接的情况。



(a) 漏斗调整器　　(b) 使用漏斗调整的 GPS 调度

图 11.7　GPS 和漏斗调整器

GPS 的延时界是这样得到的：假设每个连接上的业务流是平滑的或者说是通过了一个称为漏斗调整器的(Schwartz, 1996)。这个调整器用来限制业务的速率和突发性，它有两个参数 σ 和 ρ。一个漏斗程序使用一个可变计数器来表示速率和突发是怎样被调整的。每 $1/\rho$ s，计数器加 1，直到最大值 σ。如果一个包到达而且被发送走，那么该值减 1。如果计数器到达 0，那么就没有包可以被发送。这样，漏斗控制器就限制了包的到达速率不得超过 ρ，而一个突发中最多只能有 σ 个包被传输，也就是说一个突发的最大长度为 σ 个包。平滑的功能还可以通过一个最多拥有 σ 个令牌的令牌缓冲器(Schwartz, 1996)来实现。如图 11.7(a)所示，为一个基于令牌缓冲器的漏斗控制器。它的操作方法如下：令牌缓冲器每 $1/\rho$ s 增加 1，直到最大值 σ，每个到达的包必须拿到一个令牌才能被传送。现在让每个 GPS 连接都拥有一个漏斗调整器，拥有与其他连接相独立的参数 ρ_i 和 σ_i。得到一个如图 11.7(b)所示的 N 连接的 GPS 调度器。这里惟一的约定是漏斗控制器的速率 ρ_i 要小于等于业务的连接速率，即 $\rho_i \leqslant g_i$。可以发现，每个连接上的漏斗控制器的使用保证了连接上的最大包延时的上界，即 σ_i/g_i(Parekh 和 Gallager, 1993; Schwartz, 1996)。

① Parekh, A. K. and R. G. Gallager. 1993. "A generalized processor-sharing approach to flow control in integrated services networks: the single-node case," *IEEE/ACM Transactions on Networking*, 1, 3(June), 344–357.

如上所述的关于 GPS 公平调度算法的描述和它的特性的概括都是基于有线网络环境的。关于它在 CDMA 蜂窝无线系统中的异种业务调度上的应用和改进，已经有很多人进行了研究，其中包括 Arad 和 Leon-Garcia(1998)[①]以及 Xu 等(2002)[②]。这里介绍 Xu 等提出的 GPS 在多码 CDMA 蜂窝系统中的应用。这里只研究了上行调度，尽管很显然 GPS 也可以用于下行链路调度。Xu 等所提出的是一种基于速率的方案，每个时隙为用户动态地分配速率。每个上行用户的速率可能由于扩频因数或正交码信道的数目的不同而不同。该调度方案称为码分 GPS 公平调度方案(CDGPS)。该方案中，每个移动台在每个时隙的最后部分发送一个短信息给基站，汇报它的缓存状态(积压的包的数量)和该时隙到达的业务数量。这些信息通过随机接入信道或某个指定的上行信道发送。位于基站或更靠中心的调度器，根据从每个移动台收集来的上行信息和已知的每个移动台的 QoS 要求，计算在下一时隙应为每个移动台的上行方向所分配的传输速率，然后这个速率被转换成对应的扩频因子和使用的扩频码数目，包含在资源分配消息中通过下行信道从基站发送到移动台。

Xu 等(2002)为 CDGPS 调度机制提出的具体计算过程可以写成如下步骤。

(以下的符号 S_i，C，ϕ_i 与 GPS 机制中式(11.16)和式(11.17)的含义相同)

CDGPS：目标——计算在时隙 k 为移动台 i 分配的容量 $C_i(k)$。

令 $q_i(k-1)$代表时隙 $k-1$ 结束时，移动台 i 中积存的包的数目。令 $r_i(k)$为预计的时隙 k 中的业务到达速率(这个速率可以由第 $k-1$ 时隙到达的业务速率直接得到或进行某种平均得到)。时隙长度为 T，预计的时隙 k 后积存的包的个数为 $B_i(k)$。

按如下步骤操作：

1.
$$B_i(k) = q_i(k-1) + r_i(k)T \tag{11.18}$$

2. 参见式(11.17)，计算用户 i 可能接收到的业务量 $S_i(k)$

$$\begin{aligned} &\text{若 } B_i(k) = 0, \quad \text{则 } S_i(k) = 0 \\ &\text{若 } B_i(k) > 0, \quad \text{则 } S_i(k) = g_i T \end{aligned} \tag{11.19}$$

g_i 由式(11.16)给出。

$$\text{若} \sum_i S_i(k) < CT$$

则依据ϕ_i 分配剩余的容量。

3. 用户 i 分到的容量为：

$$C_i(k) = S_i(k)/T \tag{11.20}$$

CDGPS 作为 GPS 在 CDMA 系统下的改进版本，性能怎样？Xu 等(2002)给出了不同类型业务下的仿真和仿真结果。另外，假设使用了如图 11.7(b)的漏斗调整器，最大延时限与 GPS 得到的方法相似，仅仅多了一个调度时间 T(Xu 等，2002)

$$\text{Delay}_{\max,i} = \sigma_i/g_i + T \tag{11.21}$$

① Arad, M. A. and A. Leon-Garcia. 1998. "A generalized processor sharing approach to time scheduling in hybrid CDMA/TDMA," *Proc. IEEE Infocom' 98*, San Francisco, CA, March, 1164 - 1171.

② Xu, L. *et al*. 2002 "Dynamic bandwidth allocation with fair scheduling for WCDMA systems," *IEEE Wireless Communications*, 9, 2 (April), 26 - 32.

对于较小的时隙值,CDGPS 的最大延时限与 GPS 十分接近。

下面分析 Xu 等(2002)的仿真结果。系统的 CDMA 上行信道总容量为 $C = 2$ Mbps。时隙长度为 0.1 s。调度器不仅发送速率信息,也发送功率信息来控制小区内和小区间的干扰。假设仿真中的功率控制是理想的,传输是没有差错的,没有考虑信道衰落。仿真的第 1 个情景有 4 种不同业务,尽力而为,每个包的长度为 $L = 5120$ 比特,漏斗调整器的参数为 $\sigma = 4L, \rho = C/4$。业务流为均匀速率的泊松模型,系统负荷不断增大。最大延时限由式(11.21)得到,设 $g_i = \rho$,结果是 0.14 s,仿真得到载荷最大时的最大延时为 0.13 s,小于延时界,符合期望值。如果采用静态的分配机制,给每个业务流分配相同的固定传输速率,在载荷较重的情况下得到的延时将大大超过延时限。平均延时的结果也是类似的。在载荷较轻时,动态和静态分配技术没有太大的性能差别,如预料中的一样,因为在载荷较轻时,动态分配技术毫无用武之地。我们还进行了包含异种业务的仿真,其中业务流由 10 个语音业务、3 个视频业务和 4 个尽力而为的数据业务组成。结果显示,使用 CDGPS 调度算法得到的最大延时小于对该业务计算得到的最大延时限。每一类中的不同数据流之间的最大延时和平均延时的方差都比较小。音频和数据业务源的吞吐量的方差也较小,说明在这些业务流中得到了很好的公平性。但是,10 个语音数据流的吞吐量相差较大。

习题

11.1 在第 8 章中,知道 GSM 系统使用时隙 Aloha 随机接入机制。该过程中,上行 RACH 消息在每帧中的第 0 号时隙发送。GSM 帧长为 4.615 ms。假设只有一个频率信道用来发送 RACH 消息。

(a) 如果 GSM 的接入过程采用本章中描述的随机接入机制,那么每秒可能被发送的 RACH 消息的最大值是多少?从保证稳定性考虑,企图接入值每秒为多大是一个合理的数值?

(b) 基于呼叫阻塞考虑,在一个给定的蜂窝内的厄兰负载极值为 300 厄兰。那么,对此蜂窝内一个典型的移动终端随机接入碰撞发生的概率给出评价。在研究这个问题的过程中,对用户单位时间内的呼叫尝试和呼叫保持时间可以设置不同的值进行试验(参考第 3 章关于厄兰负载的讨论)。

11.2 在 IS-136(D-AMPS)系统下重做习题 11.1,回忆第 8 章的内容,该系统将指定频率信道中每帧中的第 1 号时隙和第 4 号时隙作为随机接入过程使用。帧长为 40 ms。

11.3 使用分析时隙 Aloha 的方法,分析没有任何时隙边界的纯 Aloha。

(a) 证明接入将在 $2T$ 时间内发生冲突。

(b) 证明纯 Aloha 的吞吐量公式为 $S = Ge^{-2G}$。证明最大可能归一化吞吐量(容量的一部分)为 $e^{-1}/2 \approx 0.18$。在此情况下重做例 11.1。

11.4 (a) 验证式(11.5)的 Gamma 分布是归一化之后的,而且它的均值是 nP_0。

(b) 将 w 代入式(11.8)得到式(11.9)的结果。

(c) 完成得到式(11.10)的推导。在式(11.10)的模型中添加一个捕获率等于 2 的条件,重新根据式(11.2)计算时隙 Aloha 的吞吐量。比较这两种结果。

11.5 参考习题 11.1。在采用了捕获效应的情况下重新计算结果。

11.6 文中给出了一系列改进随机接入的机制，其中一种采用了功率提升，另一种使用了较短的导频，还有一种提出了导频的重复。

(a) 解释这些方法能使接入性能得到改善的原因。

(b) 解释为什么 CPCH 接入机制可以提供更好的接入性能。

11.7 图 11.4 给出了两状态的语音模型。

(a) 证明处于说话状态的概率为 $\alpha/(\alpha+\beta)$，处于沉默状态的概率为 $\beta/(\alpha+\beta)$。

(b) 证明平均说话长度为 $1/\beta$，平均沉默长度为 $1/\alpha$。

(c) 在以下两种情况下计算(a)中的两种概率：1、平均说话长度为 0.4 s，平均沉默长度为 0.6 s；2、平均说话长度为 1.2 s，平均沉默长度为 1.8 s。

(d) 文中有结论说如果每个用户平均处于说话状态的概率为 0.4，那么在一个很大的系统中平均 2.5 个用户分享一个时隙。解释为什么该结论只在很大的系统中成立。

(e) 结合第 6 章中关于 CDMA 信道容量的讨论，解释为什么使用两状态语音突发模型可以获得 2.5 倍的容量增益。

11.8 (a) 按照本章所述的 PRMA 机制为语音的接入过程画一个有限状态机。

(b) 考虑文中所述的 PRMA 系统。解释为什么在每帧为 16 ms，含有 20 个时隙时，系统可以支持 56 个 PRMA 用户。为什么仿真结果显示该数据太过乐观。

11.9 针对文中所述的 PRMA ++ 接入机制。

(a) 一个用户要发起一个语音呼叫。比较使用 PRMA 机制与使用 PRMA ++ 机制有何不同。

(b) 参考图 11.5，在上行和下行方向上使用一个普通的时隙，绘出一个完整的图表，允许包有一定的延时。表示进行语音通信的多个移动台是怎样通过 PRMA ++ 机制完成各自的传输的(每个移动台设置一个序号，使用一个两状态语音模型来模拟每个移动台，注意，各移动台之间是相互独立的)。

11.10 该习题是关于文中所述的比例公平算法中两个用户例子的。用户 1 的信道质量较差，以 76.8 Kbps 或 153.6 Kbps 的速率传输，假设占空比为 50%。用户 2 信道质量较好，能以 2 倍于用户 1 的速率传输，153.6 Kbps 或 307.2 Kbps，它的占空比也为 50%。

(a) 假设基站使用轮询机制。证明：用户 1 的平均速率为 57.6 Kbps，用户 2 的平均速率为 115.7 Kbps，总的系统吞吐量为 172.8 Kbps。

(b) 用比例公平算法重复上述计算，证明：无论是单个用户的吞吐量，还是总的系统吞吐量，都有很大的改善。

11.11 对一个至少有 5 个移动台的系统，通过仿真，比较对其数据业务使用轮询调度和 PF 调度的性能差别。假设移动台可以以 3 种速率传输：76.8 Kbps，153.6 Kbps 或 307.2 Kbps。在最高速率下，一个包只需一个时隙传输。这 3 种速率的选择是基于 SIR 测量值的。假设各移动台的 SIR 的测量值每 6 个时隙变化一次，分别取 3 个可能值 10 dB，15 dB，20 dB。传输速率与这些 SIR 相对应。设计一个仿真来进行轮询机制或 PF 机制的模拟，其他的假设自行设定。进行仿真比较两种算法的(1)系统平均吞吐量；(2)单用户的平均吞吐量。

第12章　无线局域网和个域网络

前面的章节主要集中于蜂窝移动通信网络。这是原始的基于语音的方法，是传统的基于有线网络电话的扩展。第二代系统在这方面取得了极大成功，蜂窝网络的用户遍及世界各地。如第10章所述，数据流量在蜂窝网络的发展速度已经减缓，人们设计出了具有更高容量的第三代系统用于高速的分组交换。无线局域网(WLAN)随着流量数据的发展也已经取得了极大的发展，在世界上得到了广泛的应用。本章所讨论的正是这样的网络。

更确切地说，12.1节将详细讨论现在最为流行的IEEE 802.11[①]以及802.11b[②] WLAN协议，使得传输频率处于2.4 GHz，速率达到11 Mbps。同时我们将简要介绍新的有非常高传输速率的WLAN协议IEEE 802.11g[③]和802.11a[④]，这两种协议的传输频率分别是2.4 GHz和5 GHz，传输速率可达到54 Mbps。802.11g协议是802.11b协议的一个延伸，它能在相同的频率传输范围内取得更高的传输速率。12.2节将讨论个域网络，主要集中于蓝牙系统，IEEE 802.15.1协议就是针对这个应用的。IEEE 802.11b协议称为"Wi-FI"，又经常称为无线以太网。

12.1　IEEE 802.11 WLAN

局域网(LAN)是20世纪70年代晚期到80年代早期发展起来的一种连接有限距离(最长可达几千米)内的数据用户的一种有线网。光纤的普及使得局域网的商用进程取得了巨大的成就，从而能使局域网的成本控制得很低。IEEE认识到了这种网络在传输上的巨大应用价值，于是组建了IEEE 802协议委员会，负责制定一系列专门的网络标准，以解决小范围的通信问题。802系列的协议已经出现有一段时间了。这些协议主要集中于物理层与数据链路层，如同第10章所述。当前最流行的协议主要是802.3或者以太网(Ethernet)，广泛用于办公室、

---

① IEEE 802.11. IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control(MAC)and Physical Layer(PHY) Specifications, ANSI/IEEE Std 802.11, 1999 Edition, IEEE Standards Board, Piscataway, NJ.

② IEEE 802.11b. plement to IEEE Standard for Information Technology-Telecommunications and information exchange between systems-Local and metropolitan are networks-Specific requirements-Part 11: Wireless LAN Medium Access Control(MAC) and Physical Layer(PHY)specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band, IEEE Std 802.11b-1999, IEEE, January 2000.

③ IEEE 802.11g. IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control(MAC) and Physical Layer (PHY)specifications; Amendment 4: Further Data Rate Extension in the 2.4 GHz Band, IEEE Std 802.11g-2003, IEEE, New York, NY, 25 June 2003.

④ IEEE 802.11a. Supplement to IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control(MAC) and Physical Layer(PHY)specifications: High-Speed Physical Layer in the 5 GHz Band, IEEE Std 802.11a-1999, IEEE, New York, NY, 16 September 1999.

商业建筑以及校园。这个协议覆盖了数据链路的媒质控制层(MAC 层),以及下面的物理层。将要讨论的 802.11 无线协议都将覆盖上面的两种网络层次。所有的 802 WLAN 协议都是通过数据链路层中处于 MAC 子层上面的一个通用子层连接起来的,称为逻辑链路控制。将逻辑链路控制以及媒质控制层及其下面的一些子层联合起来就构成了数据链路层协议。

回顾 10.3 节所讲的每层数据单位称为那层数据的数据单元(PDU)。于是,T-PDU 称为传输层数据单元,N-PDU 称为网络层数据单元。随后,更通俗地称这两种数据单元为分组。数据链路层 PDU 以及 MAC 层的 PDU 称为帧。这些术语广泛用于 IEEE 802 协议当中,其中一些将会在本章中采用,但是会偶尔将分组看成一组数据(在第 10 章虽然试图将数据单元标准化,但是那里的字分组通常指的是其他层间的 PDU 而不是网络层的 PDU)。在本章开始的时候简要介绍一种以太网接入控制协议,然后讨论 IEEE 802.11 和 802.11b 协议。接下来介绍 IEEE 802.11g 和 802.11a 协议,这些协议都采用了第 5 章介绍的 OFDM 技术。

以太网

首先考虑有线局域网协议,IEEE 的 802.3 或者以太网协议。正如我们将要看到的,802.11 无线接入方式与早期的 802.3 接入方式十分相似,虽然由于它们处于不同的媒介上,从而造成了一定的差异。以太网利用了随机接入协议,即第 11 章讨论的 Aloha 方式。802.3 的 MAC 子层就部分采用了这种接入协议,这是一种用于用户在同一个物理链路间的连接和通信的分布式协议(这是一种简化。在这里忽略了用于 WLAN 连接的中心集线器、网桥的应用以及用于增加用户数量的虚拟 WLAN)。这种协议假设用户处于非常近的物理距离内,使得它们可以分别独立地识别出数据碰撞。因为这种识别碰撞的能力,使得以太网可以正规化它的传输容量,以超过 11.1 节介绍的 Aloha 方式的 0.368 分组每时隙(注意,这里采用的是字分组,而不是有更高正确率的字节帧)。为了使得以太网能获得高的吞吐量,协议采用了一种防碰撞机制 CSMA/CD(载波侦听多址接入/碰撞检测)。最基本的机制如下:一个希望向链路层传送"听"信号的数据基站只在检测到链路上没有能量的情况下发送数据,这是载波侦听的接入策略。显然,这样的碰撞仍然会发生,因为一个以上的采用同样接入策略的数据终端可能会同时检测到链路的空闲状态并开始传递信息,或者随后传送,使得没有足够的时间来传送到基站的数据,从而发生碰撞。在链路上的数据基站和其他的以太网基站在传送后一直检测链路状态,但是只是检测链路上的信号而不是它们要传送的信号。一旦一个基站检测到了碰撞,它就停止传送,然后延迟一个随机的时隙以减少用户的数据碰撞。如果接下来的重传仍然导致碰撞,那么重传的时隙每次都比上一次的时隙时间加长一倍。这种重传策略称为指数后退策略。注意到,为了使随机接入策略有效,基站必须靠得足够近以使得侦听其他基站的信号在传送它之前完整。这意味着物理层上传送的链路信息以时间为单位的长度比那条链路上点到点的时延扩展要长,帧在发送的时候必须能够覆盖整个链路层。时间长度取决于传输速率:越高的速率意味着越短的时间。于是,随着以太网增加它们的传输比特速率,最初能达到几十兆 bps,现在可以达到几万兆 bps,链路长度可以适当地降低了。以时间为单位进行传送的帧的长度以及以太网的长度正规化了时隙 Aloha,使得实现它成为可能。更为细节的以太网协议/802.3,包括一个简单的性能分析,可以参阅 Schwartz(1987)[①]的著作。

---

① Schwartz, M. 1987. *Telecommunication Networks: Protocols, Modeling, and Analysis*, Reading, MA, Addison-Wesley.

IEEE 802.11 接入机制

现在考虑无线以太网 IEEE 802.11 协议，也就是本章的主题。WLAN 协议包含了 802.3 协议的一些特性，如上所述。但是由于无线媒质是主要的传输载体，导致了一些大的差异。显然，无线媒质比有线媒质不可靠得多。无线扩散的特性包括时变和非对称性。媒质没有可见的传输范围；它缺乏全通性，用户基站可能无法相互间听到对方。此外，基站可能移动，导致了动态拓扑结构。这些媒质上的差异以及有线和 WLAN 用户特点的差异导致了 802.11 协议的接入策略不同于 802.3 协议。我们将讨论 802.11 协议的 MAC 层的接入策略，即分布式协同函数（DCF），这种方式带有一种特别的防碰撞机制：载波帧多址接入防碰撞（CSMA/CA），它不同于以太网中的载波侦听多址接入/碰撞检测。DCF 是一种所有移动台都能够在一个给定区字段内采用相同的接入策略进行自由通信的概念。最关键的变化是接入策略，从以太网的检测碰撞到避免碰撞。CSMA/CA 接入策略也是 MAC 层的一部分。所有 802.11 协议都采用了相同的机制。它们之间的不同在于可达的传输比特速率，以及物理层上适应不同传输速率的结构。最开始的 802.11 协议定义 1 Mbps 和 2 Mbps 的速率；Wi-Fi，也就是 802.11b 协议，采用了高达 11 Mbps 的速率；新的 802.11g 协议所设计的速率则达到了 54 Mbps。这 3 种协议全部都在 2.4 GHz 的频率上传输。802.11a 协议与 802.11g 协议近似，但是采用了 5 GHz 的频率。我们简要讨论了不同系统的物理层机制。对于传输速率非常高的 802.11a 协议和 802.11g 协议，都采用了第 5 章讨论的 OFDM 技术，以支持这种高速率的数据传输。

在讨论接入协议之前，我们简要描述了 802.11 的结构（IEEE 802.11）。IEEE 802.11 WLAN 协议定义了一个类似于蜂窝的结构，其中最基本的单元称为基本服务设施（BSS），它包含了至少两个移动台以进行相互间的通信。协议一共支持两种类型的 BSS。第 1 种是最基本的独立基本服务设施（IBSS），包括相互间的直接通信，但是没有一个管理者预先进行结构与组织分配的计划。移动台进出 BSS 以及与其他 BSS 移动台建立通信，都采用独立的方式。第 2 种是一种更为一般的方式，与本书前面所述的蜂窝系统更为接近。在这种结构里，多个 BSS 可以联合起来组成一个扩展的网络，或者可以连接成一个有线的局域网，例如以太网。在 BSS 间的无线通信或者通过电线连接的扩展网络通过一种特殊的站点连接，即接入点（AP）。在 BSS 间的移动台只用 AP 接入，于是 AP 成了从其他 BSS 或者有线网络传来信息的中继。本质上，我们可以看出 BSS 与蜂窝网络的蜂窝之间的相似之处。第 2 种 BSS 特别适合于处理移动台之间相互听不到对方消息的情况。这个方式解决了移动台之间不能直接通信的问题，也就是"隐藏终端"的问题。

IEEE 的 802.11 协议提供了两种 WLAN 结构，我们重新讨论了这个协议 MAC 层采用的载波侦听防碰撞方式。载波侦听和防碰撞意味着移动台希望传送信号以侦听媒质。如果媒质是空闲的，移动台可以传输，但是需要等待一个时间，这个时间限定为距最后一次传输的最小的时间，它用来减少另一个移动台传输信息和没有被侦听到的情况。如果媒质被侦听到忙碌，则接入将重新延迟一个特定的最小的时间，并且随机退让一段时间，在媒质被侦听到空闲后开始传输数据。这个过程降低了两个或多个移动台在同一个时间内都需要通过等待来传输信息的概率。有两种定义载波侦听的方法。第 1 种是基本的方式，也称为物理载波侦听（IEEE 802.11），在上面提到的信道上若没有任何消息被侦听，则移动台将等待一个特定的限制时间，

然后开始传送信息。第2种方法是虚拟载波侦听机制,包括在试图传送一帧数据之前交换短的保留类型的帧。对于在两个移动台之间都没有听到对方的情况下,后一种机制特别重要(这就是上面提到的"隐藏终端")。接下来,我们将详细讨论各种侦听技术。

图12.1描绘了侦听的两种基本的机制。在图12.1(a)部分,一个移动台当它准备发送消息的时候,侦听到媒质空闲;于是它等待一个特别的时间DIFS(DCI子帧空间),如果媒质仍然空闲,那么移动台开始传输数据。图12.1(b)部分描绘了媒质被侦听到忙碌的状况。移动台希望等待一个DIFS时隙后重新传输,它随机选择了一个避让的时间窗长度,以避免碰撞。然后如果在这段时间内媒质被侦听到空闲,则在这个时隙末传送数据。在试图传输之前选择等待的时间窗长度,可以降低多于一个的移动台在等待时间末尾争用信道的概率。虽然等待时间窗是随机选择的,但仍然会有碰撞发生。在这种情况下,竞争的移动台重复躲避碰撞的过程,这种方式与以太网的CSMA/CD的指数退让过程相似:一旦随机选定了一个竞争窗的长度,则在每次碰撞发生后加倍竞争窗的长度,直到达到一个最大的竞争窗长度。随后的重传在随机选取一个竞争窗长度时使用了这个最大窗长度,以便以后的重传应用。

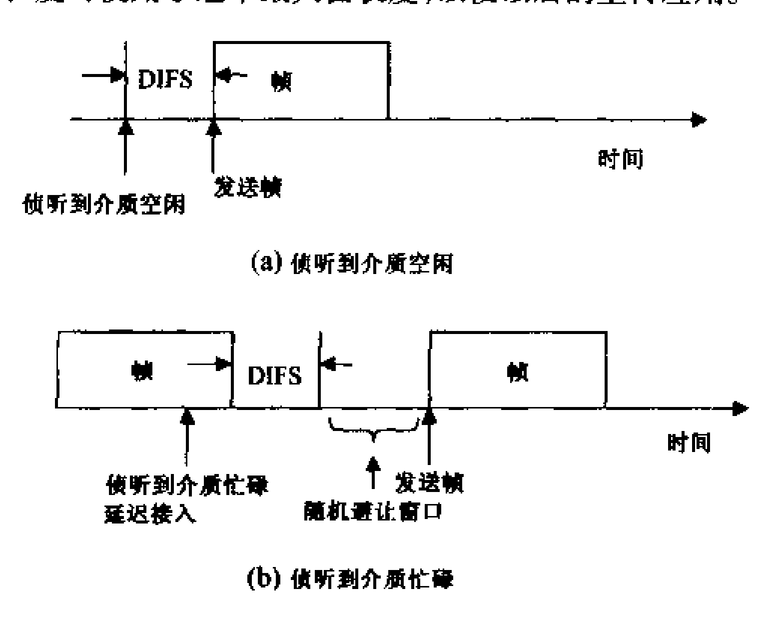


图12.1 基本接入机制

传输的每帧数据被正确地接收后,接收方会传送一个ACK确认帧。接收移动台在一个限定的时间SIFS(短帧内空间)后传送ACK帧。图12.2描述了这个过程。SIFS时隙负责接收移动台,以处理帧的接收以及帧的上层,包括MAC子层。这段时隙加上ACK帧传递到BSS的时间短于DIFS的长度,所以其他移动台在接收到ACK信号之后不能传送帧。希望传送的移动台在收到BSS间的ACK信号后必须等待DIFS一段时间(将看到虚拟机制采用的是SIFS时隙)。SIFS与DIFS时隙的使用使得一帧在传输和确认的过程中阻止了其他移动台接入媒质。

timeout信号是移动台用来确保正确接收数据帧的一种机制:一个移动台传送一帧数据设置了时间以等待ACK信号不超过一个timeout时隙。如果在timeout时隙范围以外没有接收到ACK信号,那么移动台采用一个随机避让的时间重传原始帧。802.11协议所采用的随机避让的时隙定义为许多称为时隙时间的时间单位,由达到最大数量的这样的时间单位组成竞争窗。时隙时间有效地测量移动台,检测从其他移动台传输的帧的情况,这些移动台组成了物理侦听

媒质并且声明了信道"干净"(称为信道空闲声明函数)、MAC 处理延迟、空中散布时间,以及从接收到发送的"收/发转换时间"。802.11 协议限定的时隙时间长 20 μs(SIFS 长 10 μs,DIFS 长度用 SIFS 表示为 2×时隙时间 + SIFS,即 50 μs)。移动台的竞争窗的长度在随机避让一段时间后,设定为 8 个时隙时间,随着每次重传的不成功,竞争窗的长度加倍,直到达到最大的 256 个时隙时间。随着避让时间的选取,移动台开始计数时隙时间的单位,用来决定什么时候传送。若移动台在任何一个时隙时间内侦听到信道上有传输,则马上停止计数,并且继续随机选择一个避让时间以便重传。直到等待完一个 DIFS 信号,它才重新开始对时隙时间进行计数,以确定信道空闲。

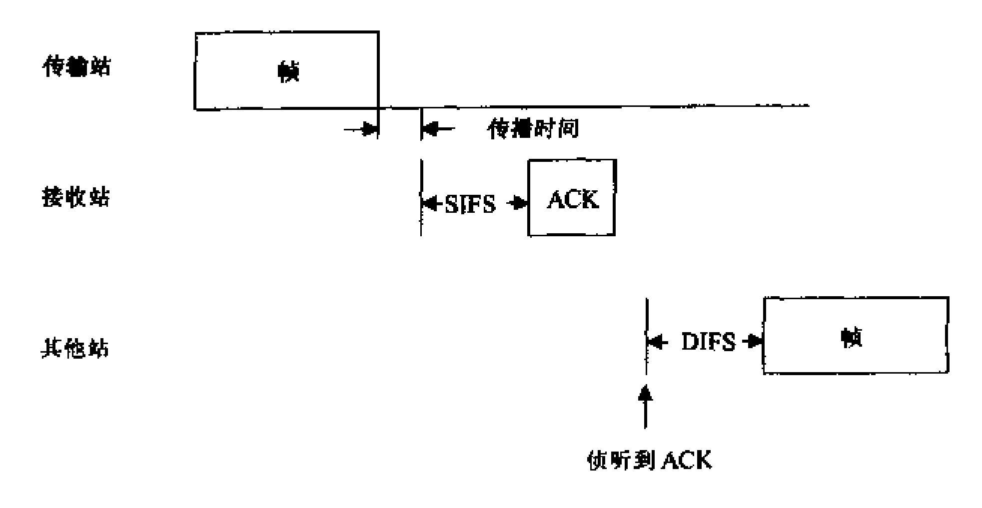


图 12.2　在基本接入中使用 SIFS/ACK 过程

注意到早期的 802.11 协议提供了两种机制用于载波侦听。第 1 种是物理载波侦听,也就是上文叙述的。第 2 种是虚拟载波侦听,是交换发送源于接收目的短帧的一种方式,警告其他移动台分组的传送将在一个特定的时间内发生。本质上是发送源希望其他用户为它保留信道以便传送。碰撞将在这段时间内避免。图 12.3(IEEE 802.11)描述了这种虚拟载波侦听机制。一个移动台希望在局域网内向一个目的地传送数据,首先得传送 RTS 信号帧。这个帧包括持续信息,发送源发送数据的长度。注意到 RTS 帧只在媒质经过 DIFS 时间段后仍然是空闲的时候才被发送,同样的过程用于基本的机制。目的移动台经过 SIFS 时间段后回复发送端一个 CTS 帧,这个帧也包括持续信息。发送源接收了 CTS 帧,在经过 SIFS 时间段后传送所需的数据帧。目的移动台又经过了 SIFS 时间段后才确认数据接收的正确性。于是,在发送源和接收目的地之间有了两次交换过程,这不同于只有一次基本接入的情形。注意到其他移动台也能听到媒质上的 RTS 信号。每个移动台采用 RTS 内的时间持续信息来更新网络分配向量(NAV)(早期的交换也许预先保留了信道)。在侦听到从目的移动台传送到发送源的 CTS 帧时,每个移动台用 CTS 内的时间持续信息来更新 NAV。注意到不是所有的移动台可以侦听到 RTS 信号帧的。侦听 CTS 信号帧提供了一种方式来更新 NAV。图 12.3 下面的 NAV 就代表了这种情形。所有的移动台而不只是源-目的移动台需要延迟一个 DIFS 时间段,以用来确认数据帧已被目的移动台正确地接收了,如图 12.3 所示。

注意到两种接入过程需要传输或者交换一系列的 MAC 帧。这些各式各样的帧包括现在所描述的数据帧。首先,所有的 802.11 MAC 帧遵循着如图 12.4(a)(IEEE 802.11)所示的帧结

构。最初是两个8位字节共16比特的帧控制字段。这个字段的分解如图12.4(b)所示。协议的版本字段指示出了协议的版本。这里版本字段被设置成0。类型9字段指示了3种帧类型：控制、数据和管理，或者特别表示的帧。在这3种类型中，精确的帧类型需要用4比特的子类型字段。本节没有讨论802.11协议的管理帧，比特数为00，控制帧为01，数据帧为10。在各种各样的控制帧里，上文提到了2种接入机制，RTS帧子帧号为1011，CTS子帧号为1100，ACK子帧号为1101。考虑到两个1比特的DS字段如图12.4(b)所示。DS意味着分布式系统，一个提供802.11局域网的灵活性的结构。分布式系统的概念允许多个BSS接入以及连接。当正常的物理距离不能提供网络需要的覆盖时，这点特别有用(IEEE 802.11)。帧控制字段中的1到达DS比特在数据帧内设为1。在其他帧内设为0。相同地，在所有数据帧内来自DS比特设为1，其他帧设为0。接下来讨论图12.4(b)内的控制帧字段。更多的片断字段在数据或者管理帧被分解为多个帧时设为1，预示着一个帧的片断接下来将要传送。重传字段用于设计先前数据帧或者管理帧的重传；更多的数据字段用于节省移动台的能量，意味着一个移动台的多个帧被缓存于AP节点上；WEP以及命令字段不在介绍的802.11协议范围内，分别用于特别的算法和服务类。

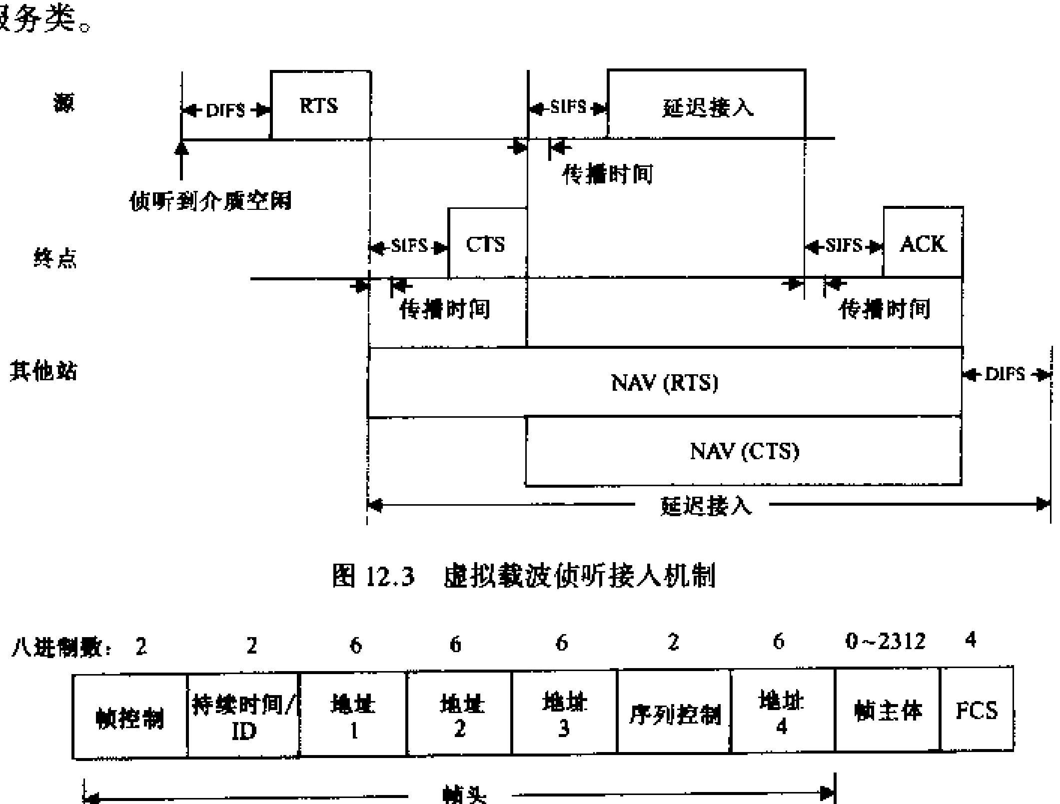
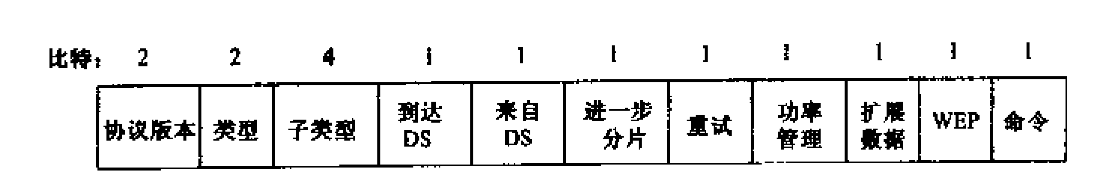


图12.3 虚拟载波侦听接入机制



八进制数：	2	2	6	6	6	2	6	0~2312	4
	帧控制	持续时间/ID	地址1	地址2	地址3	序列控制	地址4	帧主体	FCS

帧头

(a) 通用MAC帧格式

比特：	2	2	4	1	1	1	1	1	1	1	1
	协议版本	类型	子类型	到达DS	来自DS	进一步分片	重试	功率管理	扩展数据	WEP	命令

(b) 帧控制域

图12.4 MAC帧格式，802.11(引自IEEE 802.11)

现在回到图 12.4(a)的通用帧结构。我们将讨论不同用途与版本的各种各样的帧结构。所有的帧在末尾带有 4 个字节的循环冗余校验码(FCS)。这 32 比特的循环冗余校验码用于检测帧的传送是否正确,它校验所有的帧结构,包括可能存在的帧体(回忆在第 7 章中讨论的循环冗余校验码。32 个特定的生成多项式用于 802.11 帧中 FCS 字段的计算)。最后,注意到一共有 4 个地址字段,然而这些并不是所有的。

以数据帧作为第 1 个例子。802.11 协议定义了许多数据帧类型。为了简化讨论,我们采用最简单、最普通的情况,移动台被限定在一个给定的处于给定服务区字段的 BSS 内,移动台之间相互传送数据。在这种情形下,我们将上面讨论的控制帧内的到达 DS 字段和来自 DS 字段全部设成 0,如图 12.4(b)所示(其他数据类型在 IEEE 802.11 协议中有介绍)。图 12.4(b)所示的子类型字段被设定成 0000。在这些数据帧内,只采用了 3 个 6 字节的地址[图 12.4(a)所示的第 4 个字节字段没有出现]。这些分别组成了 MAC 目的地址、发送源地址以及 BSS 的 id,这个 id 是发生通信的那个单元的 id。帧体带有数据帧的信息和数据部分,即 MAC 服务数据单元(MSDU)。这个字段的最大长度为 2312 字节,如图 12.4(a)所示。MSDU 包含了 MAC 层从发送源到接收目的地的数据。如图 12.4(a)所示的 16 比特序列控制字段的数据帧有 2 部分:12 比特序列号字段,以及 4 比特分片数字字段。移动台传送的每个 MSDU 都附加了一个序列号,从 0 开始每次递增 1 直到达到最大值 4095,然后循环回 0。分片号最初被设置成 0,在这种情况下 MSDU 传输多个分片,每个分片号连续递增 1。如图 12.4(b)所示,控制帧字段的更多分片比特预示着是否有多余的 MSDU 分片将要被传输。下面介绍图 12.4(a)所示的数据帧的持续字段,数据帧类型的字段,这个字段的设计取决于更多分片比特的数据帧的类型。如果那个比特被设置成 0,传输 ACK 信号的持续值以 ms 计数,并且加上一个 SIFS 的时间段。如果这个比特被设为 1,则意味着这个数据帧更多的分片将被传输,再加上两个 ACK 帧以及 3 个 SIFS 时间段。

接下来讨论上面所描述的基本接入机制,以及虚拟接入机制的其他类型的帧结构。它们是图 12.2 和图 12.3 所示的 ACK,RTS 以及 CTS 帧。它们都是控制帧。图 12.5 所示的是它们独特的帧结构。图 12.5(a)所示的是 ACK 帧格式,它只携带一个接收机地址。回忆图 12.2 所描述的基本接入数据帧的过程。目的接收机在确认收到了正确的信息后,总是回送发射源移动台一个 ACK 信号。接收机地址字段带有发射源移动台地址,它是从接收数据帧的第 2 地址字段复制而来的发射源地址。如果数据帧在它的控制帧字段带有一个多 0 分片比特,那么 ACK 帧的持续字段设置成 0。如果数据帧的更多分片比特设置成 1,那么 ACK 持续字段所设的值比数据帧要求传送的 ACK 信号再加上 SIFS 时间段所需要的值要小。图 12.5 所示的其他两种控制帧用于完成虚拟接入过程。首先,考虑图 12.5(b)所示的 RTS 帧格式。注意到这个帧带有两个地址。第 1 个是接收机地址,即接收目的移动台地址,是这个数据帧等待被传送的地址。第 2 个地址是发射地址,即发射 RTS 的源移动台地址。持续字段值带有有用信息,如前面所述,其他移动台侦听到 RTS 帧散布以用来更新 NAV 向量,这是虚拟接入机制的部分。这个信息包含以 ms 计数的需要传送数据帧且等待发射源移动台的时间、发送一帧 CTS 的时间、一个 ACK 帧的时间以及 3 个 SIFS 的时间。以上时间段相加就是传送一个数据帧与接收到确认信息的总时间。图 12.3 是 RTS 时间。如图 12.5(c)所示的 CTS 帧格式类似于 RTS 帧,除了只使用一个地址。如图 12.5(c)所示的接收目的地址是在从发送地址字段处理 RTS 帧所恢复的部分中复制而来的。持续值也是由 RTS 帧的持续字段所携带,小于传送 CTS 帧和 SIFS 的时间。

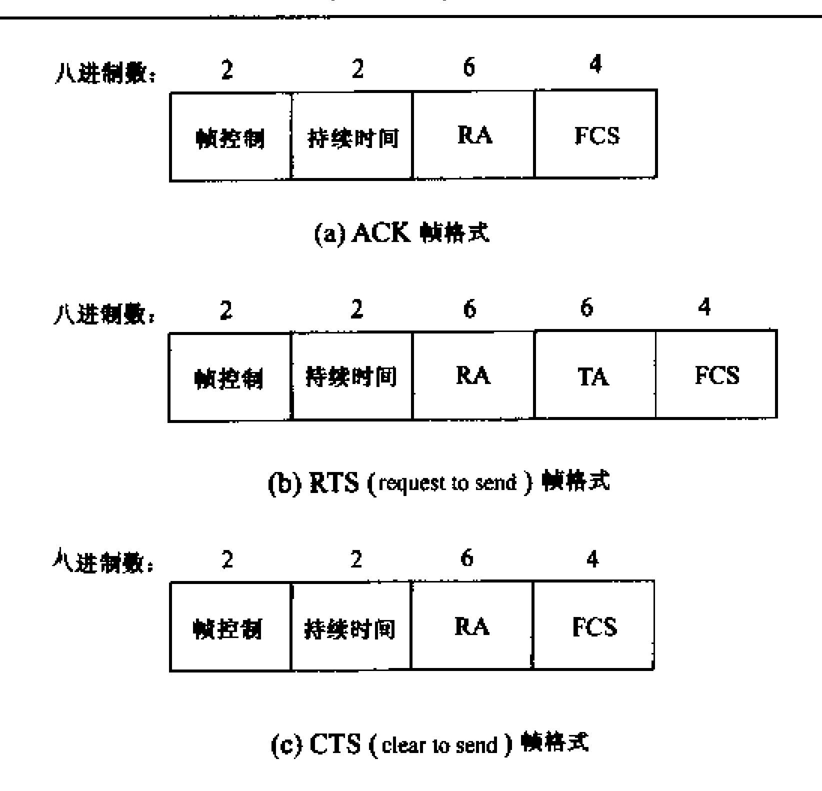


图 12.5 基本和虚拟接入的帧格式

802.11b 和 802.11 物理层结构

上面描述的是 IEEE 802.11 的接入机制。这部分简要叙述最初的 1 Mbps 和 2 Mbps 的 802.11 物理层结构,以及现在被广泛采用的传输速率在 5.5 Mbps 到 11 Mbps 的 802.11b,Wi-Fi 物理层结构。两者都采用了直接序列扩频(DSSS)编码技术以及 PSK 调制方式,在无线信道上以达到最大的传输增益。两者同时还在 2.4 GHz 的频率上进行传输。在下一节我们将简要介绍新的 802.11 协议,新协议提供了高达 54 Mbps 的速率。所有这些版本的 802.11 协议采用了相同的接入机制,即上一节所介绍的那种。有关 802.11b 协议的细节部分,建议读者查看 IEEE 802.11b 协议。最初的低比特传输速率的 802.11 协议的细节可以查阅 IEEE 802.11 协议。

物理层映射到 MAC-PDU,MAC 子层协议数据单元,或者称为 MAC 帧,是在无线传输媒质中采用的传送和接收信息的格式。所有版本的 802.11 协议,物理层帧都有一个 144 比特的前导序列,然后是 48 比特的指向正在传输的 MAC-PDU 物理层指针。IEEE 802.11b 协议一共定义了两种不同系列的物理层前导序列和指针,一个是命令,另一个是选择。在命令情形下,前导序列指针的合成部分如果以 1 Mbps 速率,差分 PSK 或 DPSK 调制方式进行传输需要 192 μs。PPDU 的数据部分的速率以 11 Mbps 或者 5.5 Mbps 进行传输。合成的 PPDU 在 802.11b 系统中称为长 PPDU 格式。在 802.11b 系统中一个短的可选择的 PPDU 格式需要花费 96 μs 传输前导序列指针的合成部分,这样可以提高 802.11b 系统的性能。

最初的 802.11 协议的 1 Mbps 或者 2 Mbps 版本,在 2.4 GHz 的频率上进行传输,采用 11 比特的巴克码,即 +1, -1, +1 +1, -1, +1 +1 +1, -1, -1, -1 来扩频正在传输的 MAC 帧的每个比特。1 Mbps 速率采用 DPSK 调制方式,2 Mbps 速率采用 DQPSK 调制方式,即差分 QPSK 方式。更高传输速率的 IEEE 802.11b 系统采用相同的 2.4 GHz 的频率进行传输,但是采用互补码键,CCK 结合 DQPSK 调制方式,从面取代上面的巴克码以获得 5.5 Mbps 和 11 Mbps 的传输

速率。在这种方式下,8 个比特同时被处理可获得 8 比特的码字。在每组中 $d_0, d_1, d_2, \cdots, d_7$ 称为连续数据比特。下面我们介绍 11 Mbps 系统。这个系统的每对连续的数据比特被分配成一组,每组分别获得(d_i, d_{i+1})4 个角频率$(0, \pi, \pi/2$ 和 $-\pi/2)$。这 4 个可能的比特对和相位值被编码成 $00 \rightarrow 0, 01 \rightarrow \pi/2, 10 \rightarrow \pi, 11 \rightarrow -\pi/2$。现在用 φ_1 代表比特对(d_0, d_1)的联合相位角,用 φ_2 代表比特对(d_2, d_3)的联合相位角,用 φ_3 代表比特对(d_4, d_5)的联合相位角,用 φ_4代表比特对(d_6, d_7)的联合相位角。这 4 个不同的相位采用 8 比特 CCK 码字 $\boldsymbol{c}$,写成复杂的向量形式,采用 11 码片/s 的扩频码

$$\boldsymbol{c} = \left[e^{j(\varphi_1+\varphi_2+\varphi_3+\varphi_4)}, e^{j(\varphi_1+\varphi_3+\varphi_4)}, e^{j(\varphi_1+\varphi_2+\varphi_4)}, -e^{j(\varphi_1+\varphi_4)}, e^{j(\varphi_1+\varphi_2+\varphi_3)}, e^{j(\varphi_1+\varphi_3)}, -e^{j(\varphi_1+\varphi_2)}, e^{j\varphi_1}\right] \tag{12.1}$$

8 比特码字工作在 1.375 MHz 的速率上,采用串行-并行方式转换接收的 11 Mbps 数据流到 8 比特码字。这个联合信号占用了与 2 Mbps 的基于巴克码的 802.11 协议相同的带宽。例如,一组 8 比特数据 00101011。令 $\varphi_1 = 0, \varphi_2 = \pi, \varphi_3 = \pi, \varphi_4 = -\pi/2$ 。相关的 $\boldsymbol{c}$ 为$[-j, j, j, j, 1, -1, 1, 1]$。

这些复数向量码字为 802.11b 协议的 11 Mbps 系统提供了高速扩频信号。5.5 Mbps 系统同样采用了式(12.1)的 8 比特码字。但是,它采用了 4 个数据比特信号、DQPSK 编码和不同于 11 Mbps 系统的 4 个相位角。IEEE 802.11b 协议的教程讨论了 CCK 编码,读者可参阅 Pearson (2001)①的著作,它包括可能的调制方式以及接收机框图。衡量 5.5 Mbps 和11 Mbps 系统的接收比特和分组误码率性能的信噪比函数,在 Heegard 等(2001)② 的著作中有介绍。假设信号处于加性高斯白色噪声环境,对比这些系统以及基于巴克码方式的低比特率系统,提出了 802.11b 协议。这个选择方式称为二进制卷积码分组或者称为 PBCC。Heegard 在文中描述了它的性能曲线,指出了 5.5 Mbps CCK 方式比 2 Mbps 巴克码系统需要 2 dB 或者更多的信噪比。11 Mbps CCK 方式比 5.5 Mbps 系统需要额外的 3 dB 信号功率。

802.11g 和 802.11a 系统的物理层结构

正如先前描述的,最新的 802.11 WLAN 协议族包括了 802.11g 和 802.11a 协议,它们都采用了正交频率复用技术以提供高达 54 Mbps 数据传输的速率。802.11g 协议在 2003 年完成标准化,并且采用与最初的 802.11 以及 802.11b 协议相同的 2.4 GHz 频率。实际上,这是与 802.11b 协议相共存的一个协议。802.11a 采用 5 GHz 的频率。两者都采用了相同的 MAC 协议方案,如最初所介绍的。关于这两种协议的细节部分,读者可以查 IEEE 802.11a 和 IEEE 802.11g 协议。注意到在所有的 802.11 协议中,原始的 802.11 以及 802.11b,802.11g,802.11a 都是放在一起介绍的,新的协议都是在原有协议的基础上进行引用或修改,并且添加个别的章节。

这两种高速率传输协议的物理层结构本质上是相同的,最重要的差异是设计频率。稍晚提出的 802.11g 协议是在 802.11a 协议的基础上进行起草的。它们都采用 20 MHz 宽的频带。OFDM 方式的子载波带宽为 312.5 kHz,在第 5 章我们已经介绍了 OFDM 方式所采用的逆快速傅里叶变换。312.5 kHz 载波带宽将 20 MHz 带宽分为 64 个子载波同时进行传输。实际上,以

① Pearson, R. 2001. "Complementary code keying made simple," Application Note AN9850.2, Intersil Americas Inc., November.

② Heegard, C. *et al*. 2001. "High-performance wireless internet," *IEEE Communications Magazine*, 39, 11(November), 64–73。

20 MHz 为中心的 16.6 MHz 带宽是 OFDM 调制所采用的实际带宽。52 个子载波均匀分布在这个带宽上，如图 12.6 所示。图中所示的数字从中心向右递增 1，向左递减 1，中心设为 0。注意到中心的 625 kHz 是空闲的。在 52 个子载波中，48 个用于传送数据；4 个用于传送导频信息。正如所看到的，OFDM 信号定义了 48 个复数，每个代表一个子载波。标准的 OFDM 数据速率支持 6 Mbps，9 Mbps，12 Mbps，18 Mbps，24 Mbps，36 Mbps，48 Mbps 以及 54 Mbps。除此之外，802.11g 标准还支持 1 Mbps，2 Mbps，5.5 Mbps 和 11 Mbps 数据速率，以及采用 DSSS 和 CCK 方式的 802.11 和 802.11b 协议所支持的速率。它也可以选择 22 Mbps 或者 33 Mbps 的数据传输速率。为了简化，只关注两种协议从 6 Mbps 到 54 Mbps 所支持的 8 种速率。

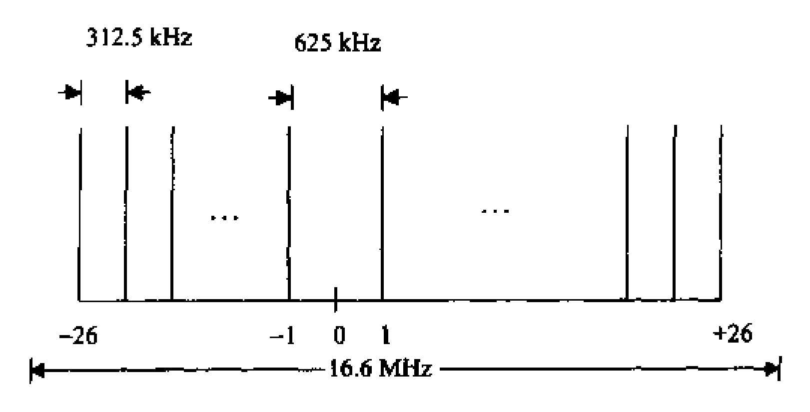


图 12.6　802.11a 和 802.11g 中 OFDM 载波分配

这 8 种速率可以通过不同的 QAM 加上 OFDM 方式获得，在第 5 章详细介绍了 OFDM 方式。特别地，6 Mbps 和 9 Mbps 采用 PSK；12 Mbps 和 18 Mbps 采用 QPSK；24 Mbps 和 36 Mbps 采用 16-QAM；最高的 48 Mbps 和 54 Mbps 采用 64-QAM，细节描述如下。

第 5 章所讨论的 OFDM 技术是正交频率技术，子载波的带宽必须是 $1/T_S$，T_S 代表 OFDM 时隙。这里子载波为 312.5 kHz，有 $T_S = 3.2\ \mu s$。从第 2 章所讨论的信道的频率选择性衰落中，我们注意到公式(2.48a)符号间隔的值在局域网中传输数据，所遇到的衰落都是非频率选择性衰落，因此在局域网中 OFDM 技术遇到的都是平衰落。这说明了为什么选择 OFDM 作为高比特传输速率的基本技术；注意到 64-QAM 系统以 54 Mbps 数据率进行传输，也就是 9 M 符号/s。如果没有 OFDM，数据传输将遇到频率选择性延迟 0.03 μs 或者更多。采用 OFDM 技术提高了数据量。快速傅里叶变换(IFFT)采用 OFDM 技术每 4 μs 进行一次。4 μs 总的时隙时间在正式标准中定义为一个符号的时间，3.2 μs 时隙为 IFFT/FFT 时隙。

图 12.7 描述了一个 802.11a 和 802.11g 采用 IFFT 算法的 OFDM 发射机的简化版本。输入比特率为上面所述的 8 种速率之一，设为 R。前向差错纠错码采用 $K = 7$，码率为 1/2 的卷积码。编码器的输出比特流穿孔，以达到所要求的输出比特率。

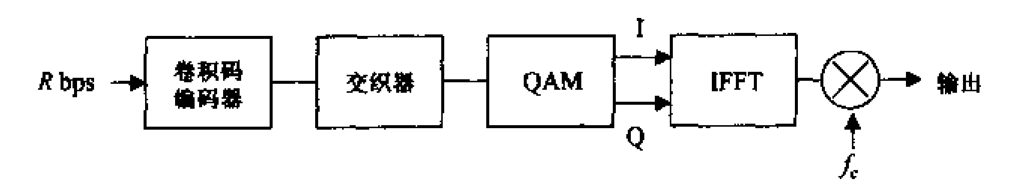


图 12.7　简化的 OFDM 发射机，IEEE 801.11a 和 801.11g

一个交织器用于穿孔输出流,存储合适数量的联合比特流提供一个 QAM 符号。每个 QAM 符号距离与 QAM 星座的复数与输入比特相关,直到存储 48 个复数以后,I 代表同相,Q 代表正交分量,每个间隔位 3.2 μs,在 IFFT 的技术器中累加。图 12.7 中的 I 和 Q 符号表示了这些复数。这 48 个相关数据被传送,以 4 为单位增加,这是因为有 4 个导频信号。采用 64 点的 IFFT,其中 16 个系数被设为 0。IFFT 输出用于系统的载波调制。

下面的两个例子描述了这个发射器的操作:

例 1

首先来看一下 6 Mbps 的情形。在这个比特速率下,每 4 μs 可以传送 24 比特信息。码率为 1/2 的编码器,在相同的时隙内可以传送 48 比特。在这种情况下没有穿孔。采用 PSK 调制方式,每 4 μs 有 48 个二进制数字,或者 +1 或者 -1,输入到 IFFT 计算器中。

例 2

现在考虑最大传输速率是 54 Mbps 的情形。这里每 4 μs 有 216 比特进入编码器。对于这个比特速率,编码器 1/3 的输出被穿孔或者丢弃,每 4 μs 的输出为 288 比特。有效编码器的速率为 3/4。在这种情况下采用 64-QAM 调制方式,采用交织编码,一次取 6 比特,用来决定 64 个可能的复数,存储后用于 IFFT 处理。注意到每 4 μs 存储 48 个这样的复数。

其他输入比特率采用相同的方式处理,所有的都需要每 4 μs 输入 48 比特到 IFFT 处理器。表 12.1 给出了 QAM 值和编码率,所有过程都要求每 4 μs 输入 IFFT 处理器 48 比特。

表 12.1 数据速率参数,IEEE 802.11a/g

数据速率(Mbps)	QAM 类型	编码速率
6	PSK	½
9	PSK	¾
12	QPSK	½
18	QPSK	¾
24	16-QAM	½
36	16-QAM	¾
48	64-QAM	⅔
54	64-QAM	¾

吞吐量性能分析

我们已经描述了 802.11 协议的物理层,以及 802.11 协议的 MAC 层的 CSMA/CA 接入机制。本节将介绍所有 802.11 结构的接入过程以及吞吐量分析。但只提供最初的 802.11 结构和 802.11b 结构的分析实例。

回忆所介绍的 MAC 层的 CSMA/CA 机制。图 12.1 和图 12.3 分别介绍了基本接入机制和虚拟接入机制。每种吞吐量性能采用无限信道容量所利用的有效性来衡量。为了回答这个问题,请参考 Bianchi①文中的分析和仿真工作,其中给出了两种接入机制的最大信道利用率,这是一个关于用户数量、最大竞争窗长度的函数。有趣的是,他采用了 p 坚持策略作为模型,与

① Bianchi, G. 2000. "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, 18, 3(March), 535 - 547.

先前章节所叙述的 PRMA 类型的结构相似。Ho 和 Chen(1996)[①]以及 Cali(1998)[②]讲述了与 802.11 系统吞吐量相关的工作。

采用 CSMA/CA 接入过程的信道利用,或者正规化吞吐量,定义了信道的部分时间,以便成功地传送数据帧的信息部分,假设每个用户移动台总有分组准备传送。决定正规化吞吐量的是吞吐量的饱和程度,即在稳定情况下系统的最大可能的吞吐量(Bianchi,2000)。为了测量这些性能,需要随机选择时隙时间和随后一段长的时间段,该时间段用来覆盖即将传输的许多数据帧以及在试图传输时发生的碰撞。长时间段的时间长度实际上用来传输被整个时隙分开的信息,这些信息的时间部分被传送,然后再决定信道利用率。相应地,跟随一个随机选择的时隙时间,其时间平均长度用来传输被平均时间分割的用户信息,这个信息提供了所需的信道利用。这个方法首先由 Bianchi 提出。用 P 和指针来定义数据帧的信息部分,包括纠错比特 H,这两者都采用了时间单位计数。一个数据帧的平均信息长度定义为 $E(P)$(在这个例子中,总是采用一个常数 P 来简化讨论)。令 T_S 是在正常传输的状态下信道被侦听为忙碌的平均时间长度。T_C 是在信道发生碰撞的状态下信道被侦听为忙碌的平均时间长度。T_S 和 T_C 都取决于特殊的接入技术,下面介绍评价 802.11 的两种接入技术的方法。采用数据帧传送时间,饱和信道利用 S 可以写成

$$S = \frac{P_{tr} P_s E(P)}{(1 - P_{tr})\sigma + P_{tr} P_s T_S + P_{tr}(1 - P_s) T_C} \tag{12.2}$$

这里,σ 是时隙长度;P_{tr}是在一个随机选取的时隙内至少有一个传输的概率;P_s 是成功传输所需的概率。注意到 $P_{tr}P_s$ 是一个移动台只在所选择的时隙内传输的概率。分子代表在一个随机选取的时间段内,传输数据帧信息部分的平均时间长度。分母代表 3 个不同的时间段,即成功传输所需信息的平均时间段。第 1 个时间段定义为空闲时间时隙;第 2 个时间段提供了没用碰撞情况下的成功传输分组所需的时间;第 3 个时间段代表着在碰撞信道下成功传输一帧数据所需的平均时间。

现在计算式(12.2)中的各项。首先考虑概率 P_{tr}。假设一共有 n 个用户移动台,令 p_d 是其中任意一个移动台在随机选择的时隙内传输的概率。这个量与第 11 章的 p 持续的概率相似。没有移动台在传输的概率是$(1-p_d)^n$。于是有 P_{tr}是在随机选择的时隙内至少有一个传输的概率,可以写成

$$P_{tr} = 1 - (1 - p_d)^n \tag{12.3}$$

为了得出成功传输的概率 P_s,采用 $P_{tr}P_s$ 表示移动台只在特定的时隙内传输。这就是 $np_d(1-p_d)^{n-1}$,即没有一个传输在$(n-1)$内。于是,得到

$$P_{tr} P_s = np_d(1 - p_d)^{n-1} \tag{12.4}$$

使用式(12.3),我们可以由上式得到 P_s。

主要的问题在于确定 p_d 在随机选取的时隙内一个给定的移动台传输的概率。为了达到这点,需要在传输前引入避让时间,这是 CSMA/CA 方式中最关键的部分。回忆先前讨论的

① Ho,T.-S. and K.-C. Chen.1996."Performance analysis of 802.11 CSMA/CA medium access control protocol,"IEEE PRMC, Taipai, Taiwan, October, 407 – 411.

② Cali, F. et al. 1998. "IEEE 802.11 wireless LAN: capacity analysis and protocol enhancement,"IEEE Infocom'98, San Francisco, CA, March.

CSMA/CA 方式,一个移动台在延迟一个 DIFS 间隔后可以侦听到媒质忙,然后以一个时隙间隔传输一个随机选择的竞争窗 w。图 12.1(b)提供了一个基本的接入机制,回忆指数避让过程,窗长度在每次碰撞发生后加倍,直到达到一个最大窗长度。于是有 $w=2^iW, 0\leqslant i\leqslant m$, m 是最大避让次数。当超过最大值时,窗长度保持为一个常数。参数 W 是最初的窗长度。Bianchi 假设数据帧以一个固定概率 p 发生碰撞,无论是第 1 次还是重传时发生碰撞(这个假设只是用来比较吞吐量以及仿真的结果)。如果碰撞概率在传输过程中相互独立,那么必须满足以下公式,p_d 是分组传输的概率。

$$p = 1-(1-p_d)^{n-1} \tag{12.5}$$

(这个公式假设至少在剩下的 $n-1$ 个移动台中有一个发生碰撞)。这就增加了在寻找概率 p 时寻找所期望的概率 p_d 的负担。但是,Bianchi 采用马尔科夫链分析,证明了两个概率之间存在更深层次的联系,如式(12.6)所示:

$$p_d = \frac{2(1-2p)}{(1-2p)(W+1)+pW(1-(2p)^m)} \tag{12.6}$$

两个非线性式(12.5)和式(12.6)可以由参量 p 和 p_d 联合解出。注意到,如果没有应用指数避让,那么 $m=0$,从式(12.6)中得出 $p_d=2/(W+1)$。这是在 $p=0$ 或 p 很小的条件下得到的 p_d 的值。我们还可以看出随着 p 的增加 p_d 单调下降。于是有 $p_d(0)=2/(W+1)>p_d(p)>p_d(1)=2/[W+1+W(2^m-1)]$。例如,如果初始窗长度为 $W>10$,单位为时隙单元,假设 p_d 小于 0.2。可用式(12.5)和式(12.6)来决定 p_d,然后可以分别计算式(12.3)和式(12.4)中的 P_{tr} 和 P_s。此外,我们已经知道了先前描述的式(12.2)的相关值 T_S 和 T_C,于是可以计算出任意两种接入机制的信道利用 S。在计算前我们可以选择合适的 p_d,使得信道利用率取最大值(Bianchi,2000)。

由式(12.2)可以看出,当 $P_s/[(T_C/\sigma)+(1-P_{tr}/P_{tr})]$ 取最大值时 S 可以取到最大值。细节部分留给读者去研究。定义 $T_C^*=T_C/\sigma$,分析显示当 $n\gg 1$ 且 $T_C^*\gg 1$ 时,使 S 取到最大值的最优的 p_d 选择可由下式估计:

$$p_{d,opt} \approx 1/n\sqrt{T_C^*/2} \tag{12.7}$$

最优的 p_d 选择仅依靠于信道在碰撞时间段内被侦听为忙碌的平均时间长度,采用时隙时间单位,也就是用户移动台在信道上传输的时间。例如,令 $T_C=1007\ \mu s$,在 11 Mbps 的 802.11b WLAN 系统中,MAC PDU 长为 1023 个字节,采用基本的接入结构。时隙时间长为 20 μs。我们可以得出 p_d, $p_{d,opt}=1/5n$ 是最优的选择。

为了进行分析,有两个变量仍然需要确认,即 T_C 和 T_S 分别是信道在碰撞时间段内被侦听为忙碌的平均时间长度,以及信道在成功传输时被侦听为忙碌的平均时间长度。这两个时间可以由前面所讨论的图 12.1 和图 12.3 中得到。特别地,基本接入机制和一个固定长度的数据分组信息字段值 P,有

$$\text{基本接入}: T_C = P + H + \text{DIFS} + \delta \tag{12.8}$$

和

$$\text{基本接入}: T_S = P + H + \text{SIFS} + \text{ACK} + \text{DIFS} + 2\delta \tag{12.9}$$

所有的单位时间量值都给出了。H 是数据帧头的长度,传输物理层信号(长模式下 192 μs)与 MAC 层帧头为 28 字节,包括 FCS 字段。δ 是散布时间,IEEE 802.11 协议定义 1 μs。我们已经知道协议定义的 SIFS 值为 10 μs,DIFS 值为 50 μs。图 12.5(a)所示的 ACK 帧格式也是如此(需

要注意的是在计算传输 ACK 帧的时间时，物理层前导和帧头的联合一定增加了帧长度）。因此需要计算在式(12.8)和式(12.9)的基本接入模式下的 T_S 和 T_C。例如，对于一个 11 Mbps 系统、1023 字节长度的 MAC-PDU，$T_S = 1221\ \mu s$，$T_C = 1007\ \mu s$。这是假设长的前导和帧头模式的情况。如果考虑到时隙时间的协议指定值为 20 μs，对于式(12.7)则有 $T_C^* = 50.3$ 和 $p_{d,opt} = 1/5n$。回忆之前的假设 $T_C \gg 1$，于是有 $n \gg 1$。所以，当 $n \geq 10$ 时，$p_{d,opt}$ 应该有效。

特别是在 $n = 10$ 的情形下的用户移动台。$p_d = 0.02$，$P_{tr} = 0.183$，$P_s = 0.895$。从式(12.2)可以得出系统的最大容量或者正规化吞吐量 $S = 0.49$。细节留给读者来考虑(注意到这些计算假设 MAC-PDU 帧长为 1023 字节。结果不同于其他数据长度的帧)。这就是说，如果数据帧可以被每个移动台获得，则 MAC-PDU 帧可以以 0.49×11 Mbps = 5.4 Mbps 的平均速率进行传输。这是一个大幅下降的 11 Mbps 吞吐量，是由于理论上的相对高侦听的失败而采用 192 μs 长前导以及帧头模式造成的(注意到 192 μs 物理层帧头在计算 T_S 时出现了两次，先于 MAC-PDU 以及 ACK 帧。即使可以忽略在解决碰撞时的平均时间 T_C，物理层帧头对于时间的影响将导致正规化容量为 $P/T_S = 0.61$，这仍然是一个偏低的值)。传送相同大小的 MAC-PDU 以及在移动台数量相同的环境下，可以获得更高的规一化吞吐量，可参考 1 Mbps 的 802.11 系统(Bianchi, 2000)。但是注意到以 1 Mbps 速率传输的数据信息，MAC-PDU，不是这里的 11 Mbps。最关键的量是 P，MAC-PDU 的长度以单位时间计数，1 Mbps 的长度为 11 Mbps 的 11 倍。物理层前导和帧头在过侦听时影响相对较小。范围在 0.8 左右的归一化吞吐量的值是可能的(Biancho, 2000)，但是这个低比特速率系统效率更高，实际吞吐量速率的量级在 800 Kbps 上下，而 5.4 Mbps 和 11 Mbps 的效率将会有一定的下降(注意到系统参数，比如在 Bianchi 文中采用的 SIFS 和 DIFS 不同于这里所采用的，但是以 1 Mbps 为速率的相对吞吐量会有很大不同。细节留给读者)。

现在考虑采用可选择的短前导序列和帧头以便减小过侦听，以 96 μs 而不是 192 μs 来传输。令 MAC-PDU 帧长一样，用户数据移动台的数量也一样。T_C 和 T_S 也分别比采用长模式的情况有所缩短，前者从 1029 μs 缩短到 $2 \times 96\ \mu s$，后者从 911 μs 缩短到 96 μs。在 11 Mbps 速率下的联合归一化容量则从 0.49 上升到 0.6。细节留给读者。一些传输效率的改进采用可选择的前导序列与帧头的联合。1 Mbps 或者 2 Mbps 传输的下界可能有所不同；192 μs 长模式已经成为高速率和低速率 802.11 系统的标准。

现在讨论虚拟接入机制情况下的归一化容量计算，仍然采用 11 Mbps 系统。式(12.3)～式(12.7)为这种情况。两者计算的差别在于基本接入机制有两个关键参数 T_C 和 T_S。回忆前些章节介绍的虚拟接入机制，如图 12.3 所总结的。采用那种描述方式，在式(12.10)与式(12.11)中需要用到两个时间(参阅 Bianchi, 2000)：

$$\text{虚拟接入}: T_C = \text{RTS} + \text{DIFS} + \delta \tag{12.10}$$

和

$$\text{虚拟接入}: T_S = \text{RTS} + 3\text{SIFS} + 4\delta + \text{DIFS} + \text{ACK} + \text{CTS} + P + H \tag{12.11}$$

注意采用虚拟接入机制传输的 MAC 层帧，MAC-PDU P 加其帧头 H，ACK 以及两个控制帧，RTS 和 CTS 在传输之前前导帧头仍然有两种传输模式，即 192 μs 的长模式与 96 μs 的短模式。在基本接入机制、11 Mbps 下，过侦听将降低系统效率。例如，在 11 Mbps 系统中传输 20 个字节的 RTS 帧需要 14.5 μs，而采用长模式需要一个 192 μs 的时间传输物理层前导序列以及帧头。

14 个字节 CTS 需要 10.2 μs。这与基本接入机制的 14 个字节的 ACK 帧计算相同。

现在考虑一个与上面基本接入机制相同的例子。仍然假设 10 个用户移动台，每个在 11 Mbps 下传输 1023 字节的数据帧。假设利用标准的长模式物理层结构。由式(12.11)和式(12.12)定义了两个时间，$T_C = 258$ μs，$T_S = 1651$ μs。注意到 T_C 是信道在碰撞时间段内被侦听为忙碌的平均时间长度，相对于先前基本接入机制的 1007 μs 已经被有效地缩短了。这样，可以将式(12.8)与式(12.10)相比较。这说明了虚拟接入机制降低了碰撞的效果。T_S 是信道在成功传输时被侦听为忙碌的平均时间长度，与先前的基本接入机制的 1221 μs 相比有所增加。这因为传输过程中需要额外的控制信息，每个都先于 192 μs 的物理层前导帧头以及随后的数据帧的传输。读者可以考虑 3 个概率用来计算在这种情况下的归一化吞吐量 $p_{d,opt} = 0.04$，$P_{tr} = 0.34$，$P_s = 0.81$。从式(12.2)可以得出 $S = 0.42$，这显然是一个低的归一化吞吐率，但是又注意到，数据帧被传递的速率为 11 Mbps。实际最优吞吐率是 0.42 × 11 Mbps 或 4.6 Mbps。与 1 Mbps 802.11 系统的结果相比较，基本接入机制提供了最大的归一化吞吐量值为 0.84(Bianchi，2000)。此外，考虑 840 Kbps 下的实际数据帧的吞吐量，可见 4.6 Mbps 的吞吐量比 802.11b 系统要小。

12.2　无线个人局域网：蓝牙/IEEE 802.15.1

在先前的章节讨论了 IEEE 802.11 WLAN 协议，用于在 100 m 甚至更远的距离实现无线高速数据传输，特别是已经商用的 802.11b 或 Wi-Fi 协议。多个这样的 WLAN 可以在比单个 WLAN 更远的距离内提供高速数据传输。个人 WLAN 将要成为下一个模式的短距离的无线数据通信(PAN)协议。它的目标是在短距离内提供无线连接，比如 10 m 范围内各种无线设备之间的连接，包括手提电脑、个人掌上电脑、数码相机，等等。蓝牙技术最初在 1998 年由数个公司提出用来达到这个目的(Bluetooth，1999[①]；Bisdikian，2001[②]；Miller 和 Bisdikian，2002[③])(蓝牙的名字来源于 10 世纪的丹麦国王 Harald Blatand，他统一了斯堪的纳维亚半岛)。(Bisdikian，2001)1999 年 3 月 IEEE 802.15 协议工作组成立，随后收录了一些蓝牙规范作为 1 Mbps WPAN 网络的 IEEE 802.15.1 协议(Bluetooth，1999，IEEE 802.15)。蓝牙/IEEE 802.15.1 协议采用与 802.11 协议相同的 2.4 GHz 频率进行传输。802.15 协议[④]的其他工作组在相同的频率内采用更高或更低的速率进行传输，包括 802.15.3 协议工作组，制定了 20 ~ 110 Mbps 的传输协议，适用于手提数据消费者的图像和多媒体业务。IEEE 802.15.4 工作组制订了更低比特的传输速率，20 ~ 250 Kbps，适用于传感器和远程控制，等等。在这节只讨论 1 Mbps 蓝牙/IEEE 802.15.1 协议。

注意到，IEEE 802.15.1 蓝牙系统在 2.4 GHz 频率上的应用。每个设备在传输时采用扩频技术，包括在 2402 ~ 2480 MHz 范围内的跳频和伪随机码。采用高斯频移键控调制，以 1 Mbps 的速率传输。典型的输出功率范围是 1 mW，使用电池供电(100 mW 以上的无线电可达到更远

---

① Bluetooth. 1999. "Specification of the Bluetooth System," Specification Volume 1, Core, v1.0.B, www.bluetooth.com, December 1.

② Bisdikian, C. 2001. "An overview of the Bluetooth wireless technology," *IEEE Communications Magazine*, 39, 12(December), 86 - 94.

③ Miller, B. A. and C. Bisdikian. 2002. *Bluetooth Revealed*, 2nd edn, Upper Saddle River, NJ, Prentice-Hall PTR.

④ IEEE 802.15. IEEE Standard for Information technology-Telecommunications and Information exchange between systems-Local and metropolitan area networks-Specific requirements, Part 15.1: Wireless Medium Access Control(MAC) and Physical Layer(PHY) Specifications for Wireless Personal Area Networks(WPANs), IEEE Std 802.15.1, IEEE-SA Standards Board, 15 April 2000.

的距离)。这种低功率模式和多个小区的联合组成了 802.11 WLAN 的模式,设备可能遇到一个功率阻碍。蓝牙设备可能覆盖 10 m 范围内组成它们的网络,每个网络由一个主机和最多 7 个从机组成,在主机的控制下从机间进行通信(Bisdikian,2001;Miller 和 Bisdikian,2002)。系统内没有竞争。网络可以相互间独立地运行并且交换空间与时间。从机和主机间也可以交换,即设备类型可以交换。任何希望接入已处于最大数量的网络中的设备需要侦听网络中的分组,在随后的时间内以一定的概率介入网络。其他的没有连入网络中但在相邻区域内的设备称为旁观者。一个单独的设备可以是多个不同又相互覆盖的网络中的设备,组成了分布网络,并且提供了比一个单独网络大的容量。但一个设备只能作为一个网络的主机。如果被连接到其他网络,它必须只能是从机。图 12.8 提供了两个相互覆盖的网络的例子。这两个网络组成了分布网络。符号 S 代表从机,M 代表主机,P 代表休眠设备,Sb 代表旁观者(Bisdikian,2001)。

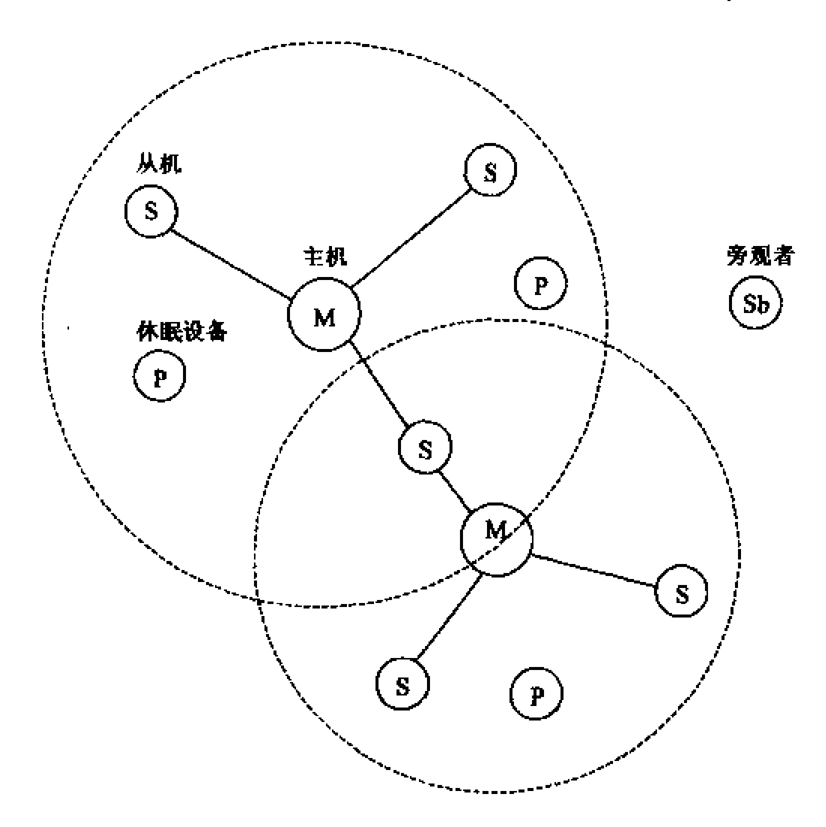


图 12.8 微微网/散射网

蓝牙规范定义成典型的层状协议。蓝牙/IEEE 802.15 协议包如图 12.9 所示(Bisdikian,2001;IEEE 802.15;Miller 和 Bisdikian,2002)。最底层是无线电层对应于 IEEE 802 协议的物理层,如图所示。这个层覆盖了整个蓝牙网络的无线电层次:1600 跳/s 的跳频扩频信号采用了 GFSK 调制方式。基带层及其上的 L2CAP 层组成了 MAC 层。蓝牙规范的传输层由物理层及 MAC 层组成,如图 12.9 所示(Bisdikian,2001;Miller 和 Bisdikian,2002)。这些协议出现在每个蓝牙规范中。MAC 层上面是中间件层,不同的蓝牙规范有不同的结构,许多只是为其他系统提供了接口。L2CAP 是逻辑链路控制和自适应协议的缩写(Logical Link Control and Adaptation Protocol),是一种特殊的蓝牙协议,提供了上层分组的接口。中间件层的 LLC 是逻辑链路控制的缩写,它允许 WLAN 以及 802 协议的网络连接到蓝牙网络中。

在网络中的成员,包括建立蓝牙通信以及在链路上传送分组,在基带中定义与规范。最基

本的通信链路由一系列长为 625 μs 的时间时隙及一系列速率为 1600 条/s 的调频序列组成。分组传输一般发生在一个时隙内，而长达 3 个或者 5 个时隙的分组传输将在下面介绍。分组传输的传输频率在 79 个频段内不断地改变。一个单独时隙分组，在分组传输时不进行跳频。在多个时隙分组传输时不进行跳频，但是在传输相同数量的单独分组时会进行跳频。时隙同步和传输频率的选择必须在主机的控制下进行。因此，蓝牙设备都带有 28 比特的时钟，每 312.5 μs 计时一次，是半个时隙单位。除此之外，每个设备带有一个惟一的 IEEE 类型 48 比特的地址。设备间的通信需要获得对方的地址。从机也需要获得主机的时钟以便同步。主机地址决定跳频序列和序列的相位，如主机时钟决定于时隙相关的频率。传送和接收信息采用时分复用，注意主机和从机在传输中可以选择：从机仅仅回应主机，主机在偶数时隙传送，从机在奇数时隙传送。

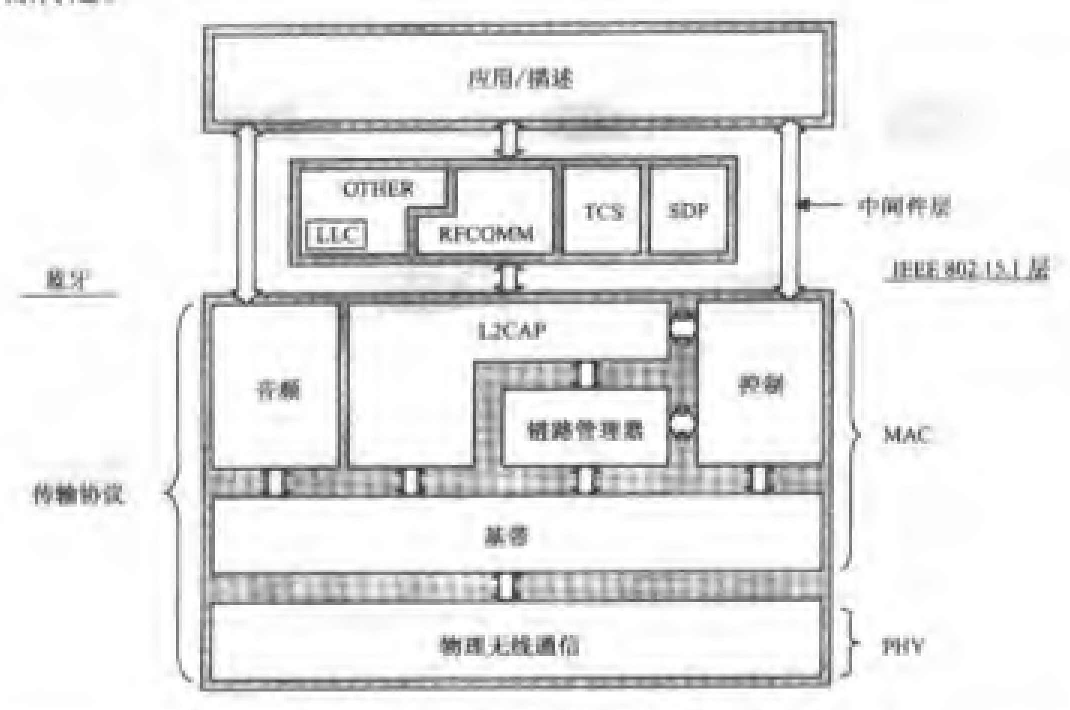


图 12.9 蓝牙 IEEE 802.15.1 协议栈(引自 IEEE 802.15, Fig.6)

微微网络以 Ad hoc 的形式连接起来，两相的查询/寻呼过程就是为了这个目的(Bisdikian, 2001)。考虑蓝牙的旁观者模式。设备上的用户控制装置用来决定是否参加网络通信或者加入邻近的网络中。利用无线电信号开始查询相位信息或者侦听页信息。下面先看一下相位询问信息。相位信息用来决定哪台设备处于传输范围内，以及可以提供的服务。发起相位信息的设备传送查询信息以邀请临近的设备加入网络。设备回应它们的地址信息。一旦与其他设备的通信建立起来，发起设备就变成为这个网络中的主机。

网络中所有的传输都由主机控制，采用时分复用的方式。主机总是发起通信的一方，向一个特定的设备传输信息，如果需要哪个设备回应主机，则它就应该回应主机(注意到主机和从机可以在任何时间交换相互的地位)。蓝牙/IEEE 802.15.1 协议支持两种类型的传输：异步最优效率数据传输和 64 Kbps 同步传输，比如语音。异步传输中，在主机和从机之间建立起异步无连接链路(ACL)。同步传输中，在主机和从机之间建立起面向连接的同步链路(SCO)。任意时间内，一个网络最多存在 3 个 SCO。ACL 链路采用分组传输、序列计数，在需要的时候使用前向纠错用来发现和纠正错误。SCO 链路不采用分组重传，但是可以使用前向纠错。

图 12.10 至图 12.12 提供了蓝牙网络主机控制的时分复用传输机制的例子[①]。一个简单的网络有 3 个设备，由 1 个主机和 2 个从机组成。图 12.10 提供了一个在基本的微微小区网络信道上的单时隙分组传输的例子。它采用了调频以及 TDD 方式进行传输。这个例子的传输由主机发起，M 代表主机，使用频率 f1 向设备 S1 传送信息。设备接收信息，然后在下一个时隙回应主机，即在 625 μs 之后。回应的分组使用频率 f2 进行传输。在下两个时隙内继续向 S1 发送以及接收从 S1 返回的信息，分别采用频率 f3 和 f4。在第 5 个时隙，主机采用频率 f5 向 S2 传输分组；在第 6 个时隙 S2 采用频率 f6 回复。正如前面所解释的，调频速率为 1600 符号/s，因此单位时隙长度为 625 μs。每次连续传输所选择的频率 f1，f2，…事先由基于主机地址的伪随机序列决定。

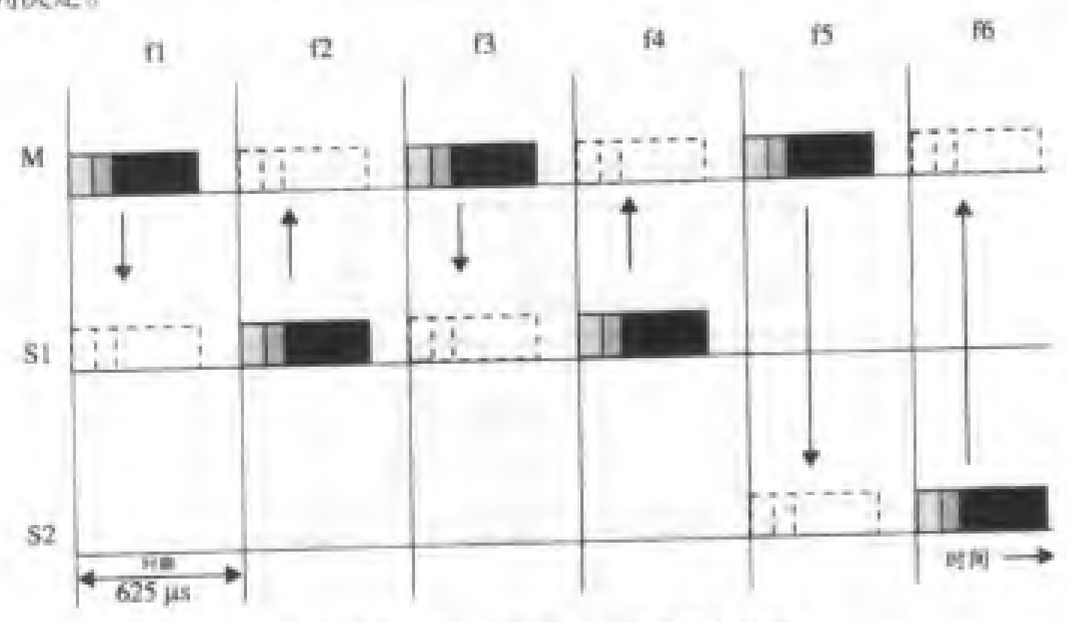


图 12.10　微微网信道，单时隙数据包传输

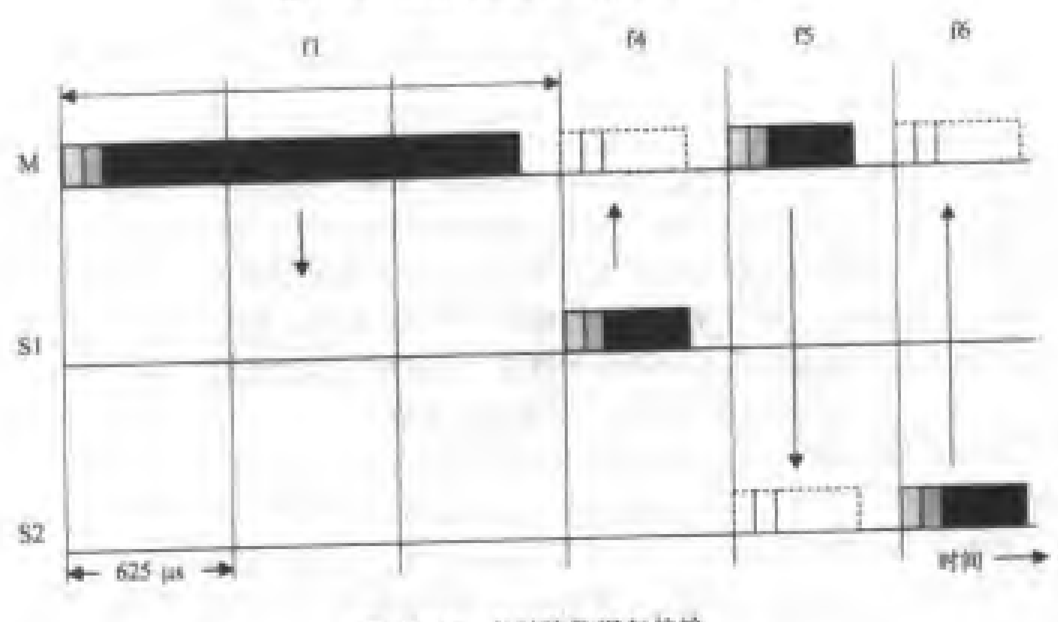


图 12.11　多时隙数据包传输

① 这些图引自 IEEE Infocom 2001 tutorial on Bluetooth，是由 IBM 的 Dr. Mahmoud Naghshineh 提出的。作者经 Dr. Naghshineh 允许使用这些图片。

图 12.11 提供了多时隙分组传输的例子。在这个例子中，主机 M 采用频率 f1 发起传输，向 S1 发送一个 3 时隙组成的分组。如果 3 个不同的分组在前 3 个时隙内，S1 的回应在时隙 4，是单时隙分组。伪随机跳频以 1600 符号/s 的速率传输。最后图 12.12 提供了 ACL 和 SCO 的联合传输模式。注意到 SCO 链路是周期性的：主机周期地选择 S1，然后 S1 在下一个时隙回应主机。ACL 链路传输用于其他的时隙。

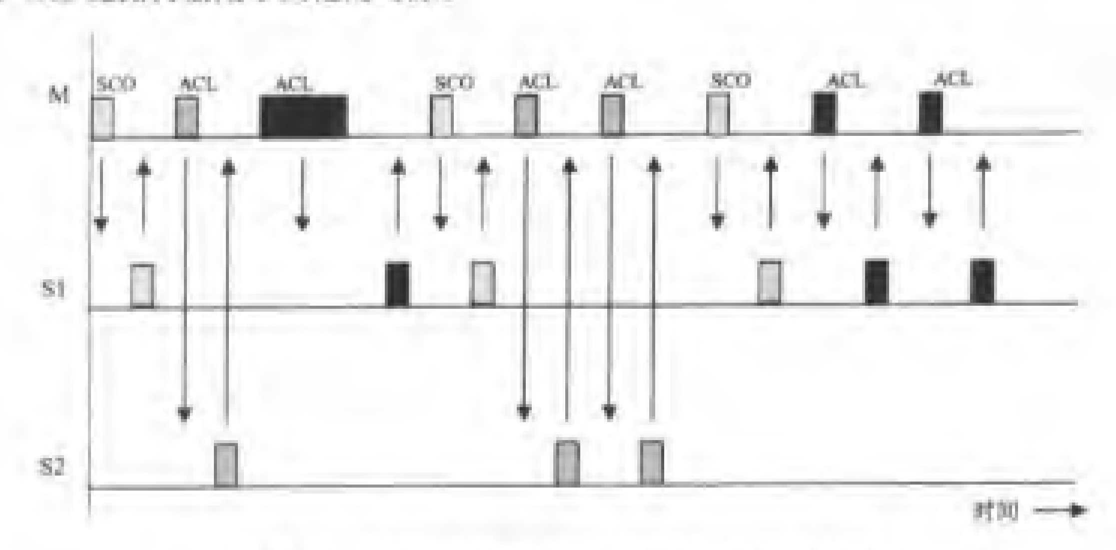


图 12.12　面向连接的同步链路(SCO)和异步无连接链路

正如所看到的，一个两相位查询/记录进程用于微微小区网络的组织。一个微微小区网络可以为发起通信的设备找到一个或者多个回应设备。例如，蓝牙设备可以寻找附近的打印机或者传真机；它可以连接到手机以及 WLAN 中，或者与 IEEE 802.15 协议的其他设备连接。如果蓝牙设备在附近没有找到其他可接入的设备，则它就会进入查询过程。在这个过程中，它周期性地发送查询消息 ID 到它临近范围内的设备。这些设备可以是先前所述的旁观者模式，如图 12.8 所示的 Sb 信号。一旦查询发起设备收集到其他设备将要进行连接的信息，它就会开启页进程以进行连接。

现在考虑收到查询信息并且准备与发起查询设备连接的旁观者模式。它仍将回应一个响应信息。查询过程如图 12.13 所示(Miller 和 Bisdikian，2002；IEEE 802.15)。注意到发起查询的设备自动成为微微小区网络主机；任何一个响应设备都是从机。查询信息的 ID 标签如图 12.13 所示，68 比特长并且带有使所有设备能够接入微微小区网络的查询接入码(一共有两种类型的查询接入码：通用查询接入码 GIAC、特殊查询接入码 DIAC)。对于所有的 802.15 特殊查询接入码都用一个共同的性质。查询信息或者 ID 信息每 625 μs 时隙间隔发送两次，即每 312.5 μs 发送一次(1 Mbps 的蓝牙系统在 68 μs 内传送完成 68 比特的分组)。查询设备，即主机等待回应信息一个时隙间隔的时间，然后传送两个或多个查询信息在下面的时隙内，重复多次这样的过程，直到链路控制认为已有足够多的设备连接上了或者达到查询的最大时间。每个分组以不同的频率传输，这些频率是从 32 个不同的频段，以伪随机方式选出的。从机一般使用与主机所使用的发送查询信息的频率不同的频率回应。查询设备选择的频率每 312.5 μs 改变一次，如图 12.13 所示。回应设备每 1.28 s 改变一次它们所听到的频率，于是侦听设备将听到查询设备的查询信息。频率的选择基于查询接入码。可以连接到微微小区网络的设备知道查询信息的频率，并且可以侦听到查询信息。图 12.13 显示了主机传送频率 f_1' 和从机回应

的频率f_i^t。上标 t 代表发起查询方，r 代表回应方。下表代表选择的特定的频率，每 312.5 μs 随机改变一次。图 12.13 显示了一个回应设备的回应信息每 625 μs 以一种特别的跳频技术选择回应频率的过程。分组包括回应设备的 48 比特的地址信息，用来识别主机设备。图 12.13 的回应设备地址用 k 表示。为了避免不同设备以相同的时间发送回应信息而发生碰撞，在接收到查询信息后设备随机选择等待一个时间后再发送回应信息。图 12.13 显示了回应设备回应第 2 个 ID 信息，在接收到查询信息后设备也可以回应第 1 个 ID 信息。回应查询信息的设备于是在主机静默或者侦听时的时隙发送回应信息。

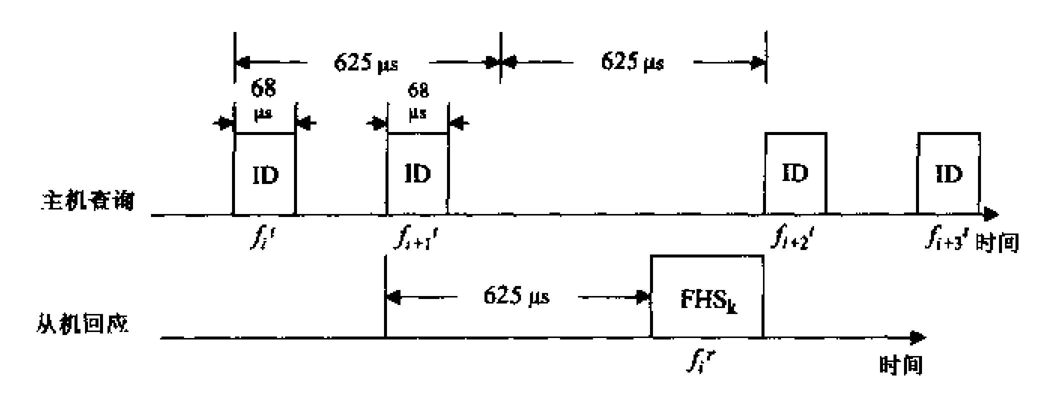


图 12.13 查询过程

注意到已经完成查询过程的主机接下来会进入记录过程来连接设备。记录过程与查询过程相似，如图 12.14 所示（IEEE 802.15；Miller 和 Bisdikian，2002）。不同之处在于主机知道所记录设备的地址。由记录设备地址决定的 32 跳频率用来传送带有设备接入码的 68 比特的 ID 分组。主机传送频率每 312.5 μs 随机改变一次，而记录设备的频率每 1.28 s 改变一次。记录设备知道回应它的设备的接入码。图 12.14 中的字母 k 代表主机记录的第 k 个设备。主机每 625 μs 在传输相位和侦听相位间选择一次。一个记录设备在一个时隙后回复自己的一个 ID 分组或者它的两个 ID 分组所带的接入码。在图 12.14 中，记录设备 k 首先以频率f_i^t 回应两个带有接入码的 ID 分组，以频率 f_i^r 传送。主机在从从机地址中接收到 ID 分组的设备接入码后，通过传送 FHS 分组确认了它传送的分组。这个分组包含从机将要确认的通信相位的 3 比特 ID，以及决定 79 个跳频频率和信道接入码的主机地址和时钟信息。从机在接收 FHS 分组时，在一个时隙后通过发送带有接入码的 ID 分组确认回复。然后，等待一个主机用来开启通信进程的特定的选择分组。这个选择分组采用图 12.10 所示的频率选择过程，在基于主机地址的 79 个伪随机跳频频率上选择频率。从机在下一个时隙时间内，从 79 个伪随机频率中选择任意一个频率回复选择分组。

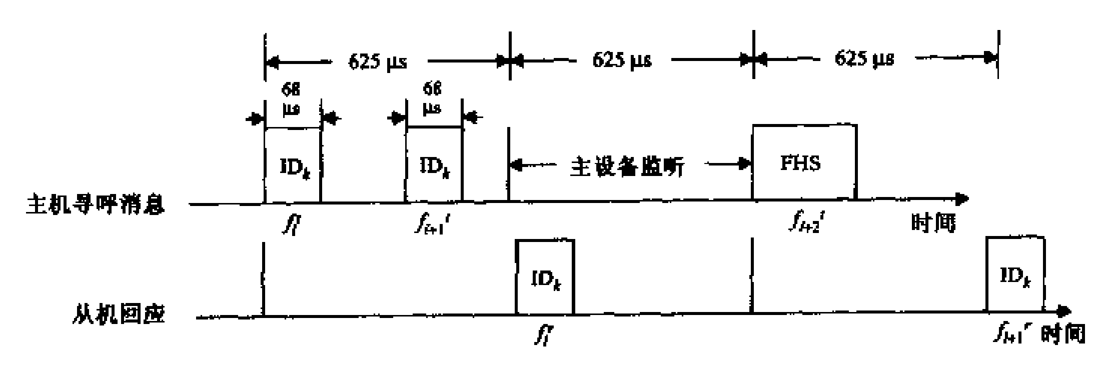


图 12.14 寻呼过程

许多基带层控制分组类型已经介绍过了。蓝牙协议中的控制分组格式以及 ACL 和 SCO 上的数据分组参照图 12.15。分组格式的开头是 72 比特的接入码(68 比特设备接入码以及上面介绍的查询接入码 ID 分组),接着传输 54 比特帧头以及有效载荷(有效载荷长度从 0 到 2745 字节,如图 12.15 所示)。接入码由 4 比特的可选 0-1 前导序列,一个由 24 比特地址演变来的 64 比特的同步字节,以及 4 比特可选 0-1 尾部组成。以上各种各样的 68 比特 ID 分组都省略了尾部,仅带有前导和同步字节。如前所述,68 比特 ID 分组用来生成同步字节的地址,它们是基于 48 比特记录设备地址中的 24 比特;用于查询设备的同步字节基于 GIAC 或者 DIAC。FHS 分组用来携带 72 比特的接入码,其中 48 比特地址用来生成同步字节。

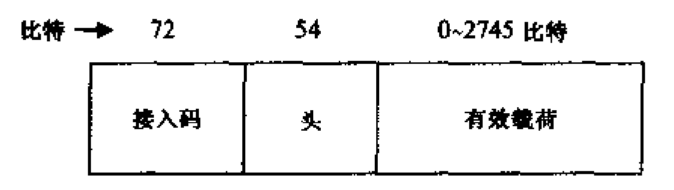


图 12.15 蓝牙数据包格式

现在考虑蓝牙分组的 54 比特帧头部分,如图 12.15 所示。帧头长 18 比特,但是每个比特被连续传输 3 次。这个重复过程是码率为 1/3 的重复码。18 比特帧头格式如图 12.16 所示。最初的3 比特,即 AM_ADDR,代表着记录过程中微微小区网络的成员 ID(回忆一个微微小区网络最多含有 7 个设备。主机使用所有的全零 AM_ADDR,而不是 FHS 分组在微微小区网络中来广播信息)。4 比特类型字段用来鉴别 16 类不同的分组类型,其中一些分组类型描述如下。3 个 1 比特标志字段,分别是 ACL 链路流控制、肯定或者否定确认分组传送,以及数据流传输的序列号。最后,8 比特的帧头错误校验字段(Header-Error-Check,HEC)是一个特殊的多项式编码,用来检测帧头的传输。所接收的不符合 HEC 校验值的整个分组将被丢弃。

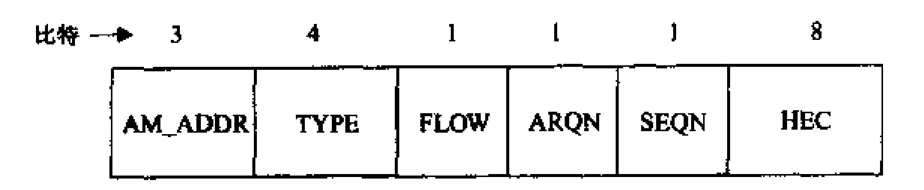


图 12.16 帧头格式

现在考虑不同蓝牙分组的传输。它们相似的结构是 4 比特的类型字段和 18 比特的帧头字段。这些分组可分为两种情况:控制分组和数据分组。控制分组是前文已经提到的 FHS 和 POLL 分组(ID 分组也是控制分组,但是它长 68 比特,所以没有在表中列出)。已经看出在查询过程中采用的 FHS 分组是微微小区网络中从机对于主机发起查询的回应,如图 12.13 所示;在记录过程中是主机告知从机发起通信的标志,如图 12.14 所示。它也用于主机与从机之间的转换。FHS 分组有效载荷包含了 144 信息比特,再加上 16 比特循环冗余校验字段用来保护信息位。这 160 比特被 80 比特采用(15,10)的汉明校验码所保护,成为 2/3 前项差错校验码(FEC)。实际的有效载荷长为 240 比特,所有的 FHS 分组长为 366 比特。注意到如果在下一个时隙需要一次应答,那么在 1 Mbps 速率的传输下将覆盖时隙时间的 59%,使得接收设备同步和接收分组。144 信息比特被分成 11 个不同的字段,包括 48 比特地址、传送分组设备的 26 比特时钟字段、3 比特从机在记录过程中的地址,以及其他类型的字段。POLL 分组不携带

有效载荷,仅由用来识别微微小区网络的 72 比特的信道接入码和 54 比特的帧头组成。正如前文所叙述的,这个分组的帧头 3 比特地址用来发起与一个特定的从机间的通信。在通信时,它也记录从机在微微小区网络中的地址。这个分组需要被选择的从机响应。一个广播的记录微微小区网络中所有从机地址的 POLL 分组被初始化为 3 比特全零地址。另一个控制分组是 NULL 分组。它类似于 POLL 分组,也不携带有效载荷,仅由 72 比特的信道接入码和 54 比特的帧头组成。这个分组用来确认主机没有发送任何查询信息或者控制流信息。

蓝牙数据分组类型有很多种。它们在 ACL 链路上进行异步数据传输,以及在 SCO 链路上进行同步数据传输。首先考虑 SCO 分组。回忆在一个微微小区网络中最多可建立不超过 3 个 SCO 链路。如果没有错误,则 SCO 分组不会重发。一共有 4 种 SCO 分组类型,主要是 64 Kbps 的语音传输。3 种 SCO 分组模式仅携带 SCO 流,类似于 FHS 分组,总共有 366 比特长,其中 240 比特是有效载荷。3 种分组类型分别是 HV1,HV2 和 HV3,以差错保护位长度递减的顺序排列。HV1 分组采用 1/3 FEC,带有有效载荷字段(如图 12.15 所示),80 个信息位以 64 Kbps 速率传播,并且重复 3 次,因此整个有效载荷位为 240 比特。这些分组可以在 SCO 链路上周期性地选择时隙来进行传输。注意到 80 信息位每隔两个时隙或者 1.25 ms 进行传输,数据传输速率为 64 Kbps。HV2 分组采用 2/3 FEC,带有 240 比特有效载荷字段,160 信息比特加上 80 比特保护位产生(15,10)汉明校验码。这些分组每 4 个时隙传输一次,对应于 64 Kbps 的语音信号。最后 HV3 分组 240 比特有效载荷位全部为信息比特,没有 FEC 保护。这些分组每隔 6 个时隙单位进行一次传输,仍然对应于 64 Kbps 的语音传输。由 3 种不同的 SCO 分组类型的效率可以由一些仿真结果得出。第 4 种分组是 DV 分组,用来携带 64 Kbps 的同步语音和数据。在一个给定的 SCO 链路上传输 DV 分组是周期性的。DV 有效载荷由 80 比特不加保护的语音字段,以及最多包含 16 比特 CRC 码的数据字段组成。CRC 码用来检测错误。我们需要采用不同的方式处理语音和数据部分,语音字段从不重复,数据字段进行错误校验,当有错误发生时进行重传。

802.15 协议定义了 7 个 ACL 分组类型。我们仅描述了其中的 6 种。所有的 6 种有效载荷字段都带有 16 比特的 CRC 校验码,没有经过 CRC 确认的需要重新传输。6 个分组类型可以分成两组:以中速进行传输或者以高速进行传输,分别以 M 和 H 表示;每个分组类型可以在 1,3,5 个时隙内传送。分别用 DM1,DM3,DM5,DH1,DH3,DH5 表示。DM1 分组的有效载荷字段最大为 17 字节(136 比特),在它之前的一个字节是帧头,它们都带有一个 16 比特的 CRC 校验码。联合的一系列比特由 2/3 的 FEC 码来保护。这些字节可以以最大速率在两个时隙内传送,最大的 DM1 数据传输速率是 108.8 Kbps。现在考虑 DH1 分组。这个分组在它的有效载荷内也带有一个字节的帧头,但是有效载荷字段的最大有效字节为 27,即 216 比特。帧头和用户有效载荷也带有 16 比特的 CRC 校验位,但是没有采用进一步的 FEC 保护。读者可以自行计算得到结果,DH1 的最大速率为 172.8 Kbps。显然,在差错率与传输速率之间有一个折中。下面考察 3 个时隙的分组。DM3 和 DH3 在有效载荷字段内最大分别可带有 121 和 183 用户有效载荷字节,占用 3 个时隙。除此以外二者都有两个字节的有效载荷帧头,并且都有 16 比特的 CRC 校验码。DM3 分组可以进一步采用 FEC 纠错码,而 DH3 方式则没有。这与单时隙的分组相同,但是,计算用户有效载荷传输的最大速率仍然非常复杂。

在计算基于 3 个分组的 ACL 用户最大数据速率时出现的两种可能的情况是:一是非对称情况,每 3 个或 4 个时隙向一个方向传输一次 DH3 或者 DM3 分组,而每 4 个时隙向另一个方

向传输一次 DH1 或者 DM1;另一个是对称情况,分组在一个方向上传输,占用 6 个时隙中的 3 个,另一个方向也是如此。计算显示非对称情况下用户最大的传输速率,它们在 DM3 和 DH3 分组一个方向上分别达到 387.2 Kbps 和 585.6 Kbps,另一个方向上分别达到 54.4 Kbps 和 86.4 Kbps。类似地,DM3 和 DH3 分组可以达到 258.1 Kbps 和 390.4 Kbps。最后,考虑到 DM5 和 DH5 分组类型。这些分组都带有 2 字节的 CRC 保护帧头和用户有效载荷字段。DM5 分组进一步可以采用 2/3 FEC,而 DH5 没有采用这个保护(注意,所有的 DM 类型的分组都采用了 2/3 FEC 保护,而 DH 类型则没有)。在这两种情况下最大用户有效载荷分别达到 224 字节和 339 字节。最大用户数据速率必须在对称和非对称两种情况下分别予以考虑。在对称情况下,DM5 分组可达到 286.7 Kbps,DH5 分组可达到 433.9 Kbps。而非对称情况下,DM5 为 477.8 Kbps/36.3 Kbps,DH5 为 723.2 Kbps/57.6 Kbps。显然,采用 DH5 类型在非对称模式下适合于在一个方向上传输的数据量非常大,而在另一个方向上数据量相对小的情况。下一节主要讨论仿真结果。非常有趣的是以不同速率传输的数据的吞吐量也不相同。

ACL 分组也用于传输数据流和来自 L2CAP 和链路管理层的控制信息,正如从图 12.9 所看到的一样,基带层支持这两个层。ACL 分组有效载荷字段采用的一个或者两个字节帧头带有一个 2 比特字段,用来区分 L2CAP 和链路管理层信息。帧头带有用户有效载荷的长度,接着是 L2CAP 和链路管理层信息。链路管理信息通常在 ACL 分组内,虽然 DV 分组也可以对它进行传输。L2CAP 层定义了一个简单的数据链路协议用来保护蓝牙基带层的高层数据源。主机-从机以及微微小区网络的操作不再出现在 L2CAP 层。这个层提供了一系列的逻辑信道,L2CAP 数据流可以是确定方向的或者不确定方向的。QoS 信息也可以在 L2CAP 层进行交换。L2CAP-PDU 也在这个层定义,它同时带有异步 L2CAP 流量,并且长度远大于 ACL 分组;L2CAP 层可以提供蓝牙协议传送的段落;接收设备相应的层也是如此。关于 L2CAP 和链路管理层可以参考 IEEE 802.15 协议以及 Miller 和 Bisdikian(2002)的著作。

蓝牙性能

有一系列的文章总结过蓝牙系统的仿真结果。我们将描述其中的一些。它们包括对于无线电网络性能的研究,以及简单地选择与用于更有效的 ACL 分组传输的分组整形机制间的比较。

无线电网络性能的研究主要集中于多个微微小区网络同时工作时的干扰(Zürbes 等,2000)[①]。这个干扰当然是以 79 个随机的跳频来衡量的。研究结果显示总体是乐观的(Zürbes 等,2000):在 10 m×20 m 的范围内,使用因特网(WWW)的数据流模型和 100 个同时传输的微微小区网络,仿真显示每 6 个 ACL 分组类型的吞吐量从 DM1 到 DH5 比预期下降不超过 6%。这些预期的吞吐量是在前文计算出的。仿真假设对数正态分布和莱斯多径衰落环境。微微小区网络 ACL 和 SCO 链路同时被仿真。ACL 链路上的 WWW 流量假设下行链路,即主机到从机;NULL 分组模型用于从机确认 ACL 数据分组的回复。WWW 会议模型由几何分布的下行数据报组成,速率为 10 个数据报每节。数据报在长度上呈对数分布,平均速率 4.1 Kbps(32.8 Kbps)以及 30 Kbps(240 Kbps)。这些数据被分割成带有用户分组类型的 ACL 分组,突发错误概率来自衰落环境下的干扰和噪声。这些错误概率正比于系统载荷,当一系列的微微小区网络同时

① Zürbes S, *et al.*, 2000. "Radio network performance of Bluetooth," IEEE International Conference on Communications, ICC, 1563-1567.

进行仿真时。例如，前向链路突发错误概率为10个微微小区网络6×10^{-3}，或者100个微微小区网络6×10^{-2}(这个系统模型的信干比确定为在非编码情况下的错误率)。有趣的是，前向链路突发错误概率的计算与ACL分组类型相同。采用长分组类型如DH5产生的干扰较小，于是越长的未经编码的分组类型对干扰越敏感。仿真测量得到的6种链路吞吐量与前文计算得到的相似。

SCO链路被模型化为指数分布，平均仿真持续时间为10 s(Zürbes等，2000)。测量SCO链路的性能采用的方法不同于ACL链路，因为SCO链路没有重传机制。相反，SCO链路采用擦除机制：接收到突发有效载荷帧头或者接入码有误就被替换成一种特殊的比特图案。突发擦除概率与帧擦除概率(FER)相等，FER是测量性能的一个重要参数；另一个是剩余比特误码率(RBER)，在SCO分组中包含没有检测出的错误。这两个性能测量正比于仿真的SCO段。结果随着SCO分组类型的不同而不同。于是，对于10段，采用HV1分组测量得到的FER是10^{-2}，HV2是0.5×10^{-2}，HV3是0.3×10^{-2}。从10段到100段，这个参数变为：HV1为0.1，HV2为0.05，HV3为0.03。剩余比特错误率显示了FEC提供的改进保护与增加干扰的效果。于是，对于100段的RBER，采用1/3 FEC的HV1测量数值为3.5×10^{-3}，采用2/3 FEC的HV2测量数值为7.5×10^{-3}，没有保护位的HV3为5.5×10^{-3}。可以看出，在HV1情况下，即使增加了干扰，FEC在减小比特错误上也有显著的效果。HV2则要承受最大的RBER，尽管不是很大，但可见2/3 FEC没有很好地克服增加的干扰。减少段的数量可以减小RBER。

前文总结的研究性能可以决定ACL数据传输策略以及语音的SCO链路性能。已经有一系列的文章介绍过蓝牙协议的MAC层。回忆在第11章介绍的无限系统中分组传输整形的关键部分。公平策略使得每个移动台有相等的机会收和发。合适的整形策略对于蓝牙系统也很关键。整形策略在蓝牙系统中可以结合主机和从机选择传送时隙的方式，主机向从机传送采用奇数时隙，而从机向主机传送采用偶数时隙。主机可以向给定的从机传送数据或者控制分组，或者只是选择那个从机而不进行数据传输，当有数据时响应数据传输。于是有一个基本的问题，主机如何在任意一个时隙内选择一个从机？这个问题显然与第11章的整形问题相似：主机应该在微微小区网络中采用什么样的整形策略？这里的问题是主机可以选择从机，当它们没有进行传输时，使得传输效率降低。最简单的整形技术是round-robin整形：主机仅循环地选择从机，它使得信道的利用低效。蓝牙中的目标是设计的整形策略可以保证公平的round-robin整形，只向需要的移动台提供服务。这个问题的解决可以在Kalia等(2000)①；Capone等(2001)②；Lee等(2002)③的著作中找到。我们将简单介绍这些文中的整形策略。

首先考虑Kalia等(2000)的文章。它假设了一个未接收的从机，所以指明当主机选择从机时设定31个比特回复位。主机可以识别那些没有收到消息的从机和根本不会有消息的从机。现在总结4种主机–从机模式：主机和从机都在等待对方发送消息，记为1-1；仅有主机向从机发送消息，从机没有消息向主机发送，记为1-0；主机没有消息发送，从机有消息向主机发送，记为0-1；主机与从机都没有消息发送记为0-0；Kalia等提出了采用这些信息的两种整形策略，并

① Kalia, M. *et al*. 2000. "Data scheduling and SAR for Bluetooth MAC," IEEE Vehicular Technology Conference, Spring, Tokyo, 716 – 720.

② Capone, A. *et al*. 2001. "Efficient polling schemes for Bluetooth picocells," IEEE International Conference on Communications, ICC2001, Helsinki, Finland, 1990 – 1994.

③ Lee, Y.-Z. *et al*. 2002. "An efficient and fair polling scheme for Bluetooth," *Proc. IEEE Milcom*, 1062 – 1068.

且把它们与 round-robin 做了比较。第 1 种策略称为优先级策略(Priority Policy,PP),在主机和从机为 1-1 时,如果主机和从机未收到对方发送的信息,则给定最高的优先级,优先级参数是一个变量。当主机和从机处于 0-0 时,主机和从机都不会发送消息,因此跳过选择。第 2 个策略称为 K 公平策略(K-Fairness Policy,KFP),在 1-1,1-0,0-1 时带有 round-robin 整形策略。运行设备在 1-1 时取时间长的服务,但是不超过 K 个时隙,K 是一个参数。仿真显示 KFP 进程提供了一个比 PP 策略更大吞吐量的方法,并且它的公平性更好。两个策略都比单独使用 round-robin 要好。

Capone 等(2001)的工作在 Kalia 等(2000)基础上进行了改进,它指出了主机未收到信息的状态并不是很精确,主机可能收不到从机指示的没有被选择的消息。他们于是提供了一种不需要从机确认消息的机制,这个机制基于前文所述的最优选择策略。这样的机制是彻底的 round-robin(exhaustive,ERR),它在一个循环的 round-robin 周期内在即将传输的设备上将所有的分组排序。这种策略显然并不如前面的公平,因为主机-从机在这里可以占用信道。为了消除这个问题并且引入公平的计量,作者提出了一种改进的办法,即限制型 round-robin(Limited Round-Robin,LRR),限制发送的分组数量。另一个改进版本是限制与加权 round-robin,(Limited 和 Weighted Round-Robin,LWRR),采用加权的方式,主要依靠观察的序列状态。一个移动台首先给定一个最大的加权值 MP。这个加权值在没有数据交换时递减 1。一旦这个加权到达最小的 1 时,它就不会重新回到 MP-1 周期而是等待很长的一段时间。任何时间数据在一个选择内交换,然后加权值再次增加至 MP。主要是比较各种不同算法的分组的延迟。仿真研究表明 ERR 策略在假设主机知道所有从机的序列长度时,彻底地服务从机以便计算出所有主机到从机的序列长度以及从机序列的长度。Kalia 的 PP 策略也提供了一个更好的策略,但它缺乏公平性。LWRR 策略在所有的数据流中表现都很好,性能接近于 ERR 并且提供了很好的公平性。两个 LWRR 改进策略可以在延迟和吞吐量性能上有更大的改进,同时提供了更好的公平性。停止选择可以减少时隙的数量而不是 MP-1 周期。这个减少显然不公平。第 2 个改变是引入了伪随机选择周期,从而减小了非公平性。如所关注的,这两个变化导致了吞吐量和延迟的改进。

习题

12.1 假设 IEEE 802.11 协议局域网环境。散布时间定为 1 μs,LAN 的尺寸是多少?

12.2 (a) 考虑一个采用 802.11 基本接入机制的数据帧传输。如图 12.4 所示,解释为什么控制字段帧设置成 0 有更多的分片比特,持续字段值代表传送一帧 ACK 加 SIFS 间隔的时间,当那个比特设成 0 时,预示着更多的分片,持续字段代表传送到下一段,加上两个 ACK 帧,再加上 3 个 SIFS 间隔的时间。文中给出了 ACK 帧的持续字段值的解释。

(b) 用你自己的话,解释虚拟接入机制的持续字段值所带有的 RTS 和 CTS 帧。

12.3 一个 BSS 带有 3 个正在相互传输数据帧的移动台。画出与图 12.2 和图 12.3 所包含的交换信息基本接入机制和虚拟接入机制相似的图。比较这两种机制。注意到一共有两个可能的情形:一个移动台随机决定用来向其他移动台传送数据的时间;一个试图回应发送移动台的接收数据帧移动台。要注意可能包括的碰撞。

12.4 考虑讨论的 802.11b 系统所采用的 8 比特码字。证明在一个给定的 00101011 的组中,相关的码字为[-j,j,j,j,1, -1,1,1]。选择一些其他的序列,重复这样的 8 比特计算。

12.5 在 IEEE 802.11g WLAN 协议中考虑为什么和有多少 OFDM 高比特速率传输可以减少码间干扰。为什么局域网环境中 OFDM 带宽可提供平衰落?

12.6 考虑表 12.1 中不同的元素。证明它们中的任意一个,在 OFDM/IFFT 计算中,每 4 μs 可以传送 48 个复数。

12.7 (a) 证明式(12.2)给出的饱和信道利用 S。特别是解释一下分母中的 3 项并且证明它们覆盖了所要成功传输一帧的时间。

(b) 证明 S 取最大时,$P_S/[(T_C/\sigma)+(1-P_{tr})/P_{tr}]$最大。

12.8 (a) 证明式(12.8)和式(12.9)的基本接入机制的参数 T_C 和 T_S。

(b) 证明式(12.10)和式(12.11)的虚拟接入机制的参数 T_C 和 T_S。

12.9 (a) 分析文中 11 Mbps 的 802.11b 饱和吞吐量基本接入机制。令 MAC-PDU 帧大小为 1023 字节。证明 10 个带有标准前导帧头的移动台的最大归一化吞吐量是 0.49。

(b) 试证帧长和用户移动台的数量,以及计算归一化吞吐量的值。给出这两个值对吞吐量的影响。

12.10 对于虚拟接入信道,重复习题 12.9,然后与基本接入信道进行比较。

12.11 图 12.10 和图 12.11 给出了蓝牙主机-从机的单时隙和多时隙分组传输的异步传输。给出了这两个例子所包含 3 个移动台的计时数据报。令多分组中分组长度在 1 个时隙到 3 个时隙间变化。图 12.10 和图 12.11 包括了使用频率的信息。

12.12 (a) 扩展图 12.13 的 3 个从机加入微微小区网络的蓝牙询问机制。

(b) 扩展图 12.14 的 3 个从机加入微微小区网络的蓝牙记录机制。

12.13 文中描述了 68 比特 ID 分组和 366 比特 FHS 分组。给出图 12.15 扩展形式的完整格式,注意长度以及每个字段和子字段的内容。

12.14 (a) 考虑一个 SCO 链路上的同步数据传输。采用一个主机和一个从机的计时数据报来证明每次传输有 3 个 SCO 分组类型。

(b) 图示 3 个不同类型的 SCO 分组数据包的有效载荷字段。

(c) 解释为什么传输这些分组类型采用 64 Kbps 的速率。

12.15 证明最大的 DM1 用户数据传输速率是 108.8 Kbps,而 DH1 是 172.8 Kbps。

12.16 考虑 3 个 ACL 分组传输。

(a) 在主机和从机间的数据报传输采用了同步和异步结合的传输方式。

(b) 证明文中各种同步和异步的在两个方向上最大的用户数据传输速率。解释为什么在异步下,DM1 和 DH1 传输速率是习题 12.15 的一半。

12.17 给出文中所描述的每 6 个 ACL 数据分组的数据报的有效载荷字段。

12.18 证明在文中异步和同步传输模式下的每 5 个时隙 ACL 分组的最大用户数据传输速率。

12.19 参考文中的蓝牙流量整形技术。一开始是一个简单的 round-robin 选择,然后提高了复杂性以改进吞吐量、延迟,以及维护公平性。

(a) 用你自己的话描述各种不同的技术,采用主机-从机计时数据报来解释可能的进程。公平的含义是什么? 为什么一项技术能够比另一项更公平?

(b) 注意到文中的解释,一共有 4 种不同情形的蓝牙流量整形技术,包括一个主机或者一个从机或者两者都等待一个分组,以及都只有空缓冲的情况。讨论采用上述 4 种情形的 round-robin 选择和一个简单的整形策略。

参考文献

Abramson, N. 1973. "The Aloha system," Ch. 14, in *Computer Networks*, ed. N. Abramson and F. Kuo, Englewood Cliffs, NJ, Prentice-Hall.

Abramson, N. 1977. "The throughput of packet broadcasting channels," *IEEE Transactions on Communications*, COM-25, 1 (January), 117–128.

Aein, J. M. 1973. "Power balancing in systems employing frequency reuse," *Comsat Technology Review*, 3, 2 (Fall).

Akyildiz, I. F. and J. S. M. Ho. 1995. "Dynamic mobile user location update for wireless PCS networks," *Wireless Networks*, 1, 2 (July), 187–196.

Akyildiz, I. F. *et al.* 1998. "Mobility management in current and future communication networks," *IEEE Network*, 12, 4 (July/August), 39–49.

Anasti, G. *et al.* 1996. "A bandwidth reservation protocol for speech/data integration in TDMA-based advanced mobile systems," *Proc. IEEE Infocom'96*, San Francisco, CA, March, 722–729.

Anasti, G. *et al.* 1997. "A contention/reservation access protocol for speech and data integration in TDMA-based advanced mobile systems," *Mobile Networks and Applications*, 2, 3–18.

Andrews, M. *et al.* 2001. "Providing quality of service over a shared wireless link," *IEEE Communications Magazine*, 39, 2 (February), 150–153.

Arad, M. A. and A. Leon-Garcia. 1998. "A generalized processor sharing approach to time scheduling in hybrid CDMA/TDMA," *Proc. IEEE Infocom'98*, San Francisco, CA, March, 1164–1171.

Arbak, J. C. and W. Van Blitterwijk. 1987. "Capacity of slotted Aloha in Rayleigh fading channels," *IEEE Journal on Selected Areas in Communications*, SAC-5, 2 (February), 261–269.

Atal, B. S. and S. L. Hanauer. 1971. "Speech analysis and synthesis by linear prediction of the speech wave," *Journal of the Acoustic Society of America*, 50, 2 (August), 637–655.

Bar-Noy, A. *et al.* 1995. "Mobile users: to update or not to update?" *Wireless Networks*, 1, 2 (July), 175–185.

Bello, P. A. 1963. "Characterization of randomly time-variant linear channels," *IEEE Transactions on Communication Systems*, 12, CS-11 (December), 360–393.

Bender, P. *et al.* 2000. "CDMA/HDR: a bandwidth-efficient high-speed service for nomadic users," *IEEE Communications Magazine*, 38, 7 (July), 70–77.

Berrou, C., A. Glavieux, and P. Thitimajshima. 1993. "Near Shannon-limit error-correcting coding and decoding: turbo codes," IEEE International Conference on Communications, ICC93, Geneva, Switzerland, May, 1064–1070.

Berrou, C. and A. Glavieux. 1996. "Near optimum error correcting coding and decoding: turbo codes," *IEEE Transactions on Communications*, 44, 10 (October), 1261–1271.

Bettstetter, C. *et al.* 1999. "GSM Phase 2+ general packet radio service GPRS: architecture, protocols, and air interface," *IEEE Communication Surveys*, 3rd 8tr.

Bianchi, G. 2000. "Performance analysis of the IEEE 802.11 distributed coordination function," *IEEE Journal on Selected Areas in Communications*, 18, 3 (March), 535–547.

Bisdikian, C. 2001. "An overview of the Bluetooth wireless technology," *IEEE Communications Magazine*, 39, 12 (December), 86–94.

Black, D. M. and D. O. Reudink. 1972. "Some characteristics of radio propagation at 800 MHz in the Philadelphia area," *IEEE Transactions on Vehicular Technology*, 21 (May), 45–51.

Blom, J. and F. Gunnarson. 1998. "Power control in cellular systems, Linköping studies in science

and technology," Thesis No. 706, Department of Electrical Engineering, Linköpings Universitet, Linköping, Sweden.

Bluetooth. 1999. "Specification of the Bluetooth System," Specification Volume 1, Core, v1.0.B, www.bluetooth.com, December 1.

Brand, A. E. and A. H. Aghvami. 1996. "Performance of a joint CDMA/PRMA protocol for mixed voice/data transmission for third generation mobile communications," *IEEE Journal on Selected Areas in Communications*, 14, 9 (December), 1698–1707.

Brenner, D. G. 1959. "Linear diversity combining techniques," *Proc. IRE*, 47, 6 (June), 1075–1102.

Le Bris, L. and W. Robison. 1999. "Dynamic channel assignment in GSM networks," *Proc. VTC'99*, IEEE Vehicular Technology Conference, Amsterdam, September, 2339–2342.

Cai, J. and D. J. Goodman. 1997. "General packet radio service in GSM," *IEEE Communications Magazine*, 35, 10 (October), 122–131.

Cali, F. *et al.* 1998. "IEEE 802.11 wireless LAN: capacity analysis and protocol enhancement," IEEE Infocom'98, San Francisco, CA, March.

Capone, A. *et al.* 2001. "Efficient polling schemes for Bluetooth picocells," IEEE International Conference on Communications, ICC2001, Helsinki, Finland, 1990–1994.

cdma 2001. cdma2000 high rate packet data air interface specification, 3GPP2 C.S0024, Version 3.0, 3rd Generation Partnership Project 2, 3GPP2, December, 5, http://www.3gpp2.org/public_html/specs/

cdma 2002. Physical layer standard for ccdma2000 spread spectrum systems, Release A, 3GPP2 C.S0002-A, Version 6.0, 3rd Generation Partnership Project 2, 3GPP2, February, http://www.3gpp2.org/public_html/specs/

Chao, C. and W. Chen. 1997. "Connection admission control for mobile multiple-class personal communication networks," *IEEE Journal on Selected Areas in Comunications*, 15, 10 (October), 1618–1626.

Chen, J.-H. *et al.* 1992. "A low-delay CELP coder for the CCITT 16 kb/s Speech Coding Standard," *IEEE Journal on Selected Areas in Communications*, 10, 5 (June), 830–849.

Chih-Lin I and K. Sabnani. 1995a. "Variable speading gain CDMA with adaptive control for true packet switching wireless network," IEEE International Conference on Communications, ICC95, Seattle, WA, June, 725–730.

Chih-Lin I and K. Sabnani. 1995b. Variable spreading gain CDMA with adaptive control for integrated traffic in wireless networks," IEEE 45th Vehicular Technology Conference, VTC95, Chicago, IL, July, 794–798.

Chih-Lin I and R. D. Gitlin. 1995. "Multi-code CDMA wireless personal communication networks," IEEE International Conference on Communications, ICC95, Seattle, WA, June, 1060–1064.

Choi, S. and K. G. Shin. 1998. "Predictive and adaptive bandwidth reservation for handoffs in QoS-sensitive cellular networks," Proc. SIGCOMM98, 155–166.

Cox, D. C. and D. O. Reudink. 1972. "Dynamic channel assignment in two-dimensional large-scale mobile communication systems," *Bell System Technology Journal*, 51, 1611–1672.

Cox, D. C. and D. O. Reudink. 1973. "Increasing channel occupancy in large-scale mobile radio systems: dynamic channel assignment," *IEEE Transactions on Vehicular Technology*, VT-22, 218–222.

Cox, D. C. 1995. "Wireless personal communications: what is it?," *IEEE Personal Communications*, 2, 2, 20–35.

Dahlman, E. and K. Jamal. 1996. "Wideband services in a DS-CDMA based FPLMTS system," *IEEE Vehicular Technology Conference*, VTC96, 1656–1660.

DeVille, J. M. 1993. "A reservation-based multiple access scheme for a future universal telecommunications system," *Mobile and Personal Communications*, IEE Conference Publication

No. 387, 13–15 December, 210–215.

Epstein, B. and M. Schwartz. 1998. QoS-Based Predictive Admission Control for Multi-Media Traffic, in *Broadband Wireless Communications*, ed. M. Luise and S. Pupolin, Berlin, Springer-Verlag, pp. 213–224.

Epstein, B. and M. Schwartz. 2000. "Predictive QoS-based admission control for multiclass traffic in cellular wireless networks," *IEEE Journal on Selected Areas in Communications*, 18, 3 (March), 523–534.

Egli, J. J. 1957. "Radio propagation above 40 Mc/s over irregular terrain," *Proc. IRE*, October, 1383–1391.

Everitt, D. and D. Manfield. 1989. "Performance analysis of cellular mobile communication systems with dynamic channel assignment," *IEEE Journal on Selected Areas in Communications*, 7, 8 (October), 1172–1180.

Fattah, H. and C. Leung. 2002. "An overview of scheduling algorithms in wireless multimedia networks," *IEEE Wireless Communications*, 9, 5 (October), 76–83.

Forney, G. D. 1966. *Concatenated Codes*, Cambridge, MA, MIT Press.

Foschini, G. J. 1996. "Layered space-time architecture for wireless communication in a fading environment when using multiple antennas," *Bell Laboratories Technical Journal*, 1, 2, 41–59.

Foschini, G. J. and Z. Miljanic. 1993. "A simple distributed autonomous power control algorithm and its convergence," *IEEE Transactions on Vehicular Technology*, 42, 4 (November), 641–646.

Frenkiel, R. 2002. "A brief history of mobile communications," *IEEE Vehicular Technology Society News*, May, 4–7.

Furuskar, A. *et al.* 1999a. "EDGE: enhanced data rates for GSM and TDMA/136 evolution," *IEEE Personal Communications*, 6, 3 (June), 56–66.

Furuskar, A. *et al.* 1999b. "Capacity evaluation of the EDGE concept for enhanced data rates in GSM and TDMA/136," *Proc. IEEE Vehicular Technology Conference*, VTC99, 1648–1652.

Garg, V. K. 2000. *IS-95 CDMA and cdma2000*, Upper Saddle River, NJ, Prentice-Hall PTR.

Garg, V. K. and J. E. Wilkes. 1999. *Principles and Applications of GSM*, Upper Saddle River, NJ, Prentice-Hall PTR.

Gibson, J. D. ed. 1996. *The Mobile Communication Handbook*, Boca Raton, FL, CRC Press/IEEE Press.

Gilhausen, K. H. *et al.* 1991. "On the capacity of a cellular CDMA system," *IEEE Transactions on Vehicular Technology*, 40, 2 (May), 303–312.

Goodman, D. J. 1997. *Wireless Personal Communication Systems*, Reading, MA, Addison-Wesley.

Goodman, D. J. and A. A. M. Saleh. 1987. "The near-far effect in local Aloha radio communications," *IEEE Transactions on Vehicular Technology*, VT-36, 1 (February), 19–27.

Goodman, D. J. and S. X. Wei. 1991. "Efficiency of packet reservation access," *IEEE Transactions on Vehicular Technology*, 40, 1 (February), 170–176.

Goodman, D. J. *et al.* 1989. "Packet reservation multiple access for local wireless communication," *IEEE Transactions on Communications*, 37, 8 (August), 885–890.

GPRS. 1998. Digital cellular communications system (Phase 2+): general packet radio service description, Stage 1, EN 301 113 v6.1.1 (1998–11) (GSM 02.60 version 6.1.1 Release 1997), European Telecommunications Institute (ETSI), Sophia Antipoli, Valbonne, France.

Grandhi, S. A. *et al.* 1994. "Distributed power control in cellular radio systems," *IEEE Transactions on Communications*, 42, 2/3/4 (February–April), 226–228.

Grillo, D. *et al.* 1993. "A performance analysis of PRMA considering speech, data, co-channel interference and ARQ error recovery," *Mobile and Personal Communications*, IEEE Conference Publication No. 387, December 13–15, 161–171.

GSM World, www.gsmworld.com

Gudmundson, B. *et al.* 1992. "A comparison of CDMA and TDMA systems," *Proc. IEEE Vehicular Technology Conference*, Denver, CO, May, 732–735.

Guerin, R. A. 1987. "Channel occupancy time distribution in cellular radio systems," *IEEE Transactions on Vehicular Technology*, VT-35, 3 (August), 89–99.

Hall, E. K. and S. G. Wilson. 1998. "Design and analysis of turbo codes on Rayleigh fading channels," *IEEE Journal on Selected Areas in Communications*, 16, 2 (February), 160–174.

Heegard, C. *et al.* 2001. "High-performance wireless internet," *IEEE Communications Magazine*, 39, 11 (November), 64–73.

Heller, J. A. and I. M. Jacobs. 1971. "Viterbi decoding for space and satellite communications," *IEEE Transactions on Communication Technology*, COM-19, 5, part II (October), 835–848.

Hong, D. and S. S. Rappaport. 1999. "Traffic model and performance analysis for cellular mobile radio telephone systems with prioritized and non-prioritized handoff procedures," *IEEE Transactions on Vehicular Technology*, VT-35, 3 (August), 77–92. See also CEAS Technical Report No. 773, 1 June 1999, College of Engineering and Applied Sciences, State University of New York, Stony Brook, NY 11794.

Ho, T.-S. and K.-C. Chen. 1996. "Performance analysis of 802.11 CSMA/CA medium access control protocol," IEEE PRMC, Taipai, Taiwan, October, 407–411.

IEEE. 1999. "The evolution of TDMA to 3G," *IEEE Personal Communications*, 6, 3 (June).

IEEE 802.11. IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) Specifications, ANSI/IEEE Std 802.11, 1999 Edition, IEEE Standards Board, Piscataway, NJ.

IEEE 802.11a. Supplement to IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: High-Speed Physical Layer in the 5 GHz Band, IEEE Std 802.11a-1999, IEEE, New York, NY, 16 September 1999.

IEEE 802.11b. Supplement to IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications: Higher-Speed Physical Layer Extension in the 2.4 GHz Band, IEEE Std 802.11b-1999, IEEE, January 2000.

IEEE 802.11g. IEEE Standard for Information technology-Telecommunications and information exchange between systems-Local and metropolitan area networks-Specific requirements-Part 11: Wireless LAN Medium Access Control (MAC) and Physical Layer (PHY) specifications; Amendment 4: Further Data Rate Extension in the 2.4 GHz Band, IEEE Std 802.11g-2003, IEEE, New York, NY, 25 June 2003.

IEEE 802.15. IEEE Standard for Information technology-Telecommunications and Information exchange between systems-Local and metropolitan area networks-Specific requirements, Part 15.1: Wireless Medium Access Control (MAC) and Physical Layer (PHY) Specifications for Wireless Personal Area Networks (WPANs), IEEE Std 802.15.1, IEEE-SA Standards Board, 15 April, 2000.

Jabbari, B. 1996. "Teletraffic aspects of evolving and next-generation wireless communication networks," *IEEE Personal Communications*, 3, 6 (December), 4–9.

Jain, R. and Y.-B. Lin. 1995. "An auxiliary location strategy employing forwarding pointers to reduce network impacts of PCS," *Wireless Networks*, 1, 2 (July), 197–210.

Jakes, W. C. ed. 1974. *Microwave Mobile Communications*, AT&T, 1995 edition, New York, IEEE Press.

Jayant, N. 1992. "Signal compression: technology targets and research directions," *IEEE Journal on Selected Areas in Communications*, 10, 5 (June), 796–818.

Kalden, R. *et al.* 2000. "Wireless internet access based on GPRS," *IEEE Personal Communications*, 7, 2 (April), 8–18.

Kalia, M. *et al.* 2000. "Data scheduling and SAR for Bluetooth MAC," IEEE Vehicular Technology Conference, Spring, Tokyo, 716–720.

Karol, M. J. *et al.* 1995a. "An efficient demand-assignment multiple-access protocol for wireless packet (ATM) networks," *Wireless Networks*, 1, 5 (October), 267–279.

Karol, M. J. *et al.* 1995b. "Distributed-queueing request update multiple access (DQRUMA) for wireless packet (ATM) networks," IEEE International Conference on Communications, ICC'95, Seattle, WA, June, 1224–1231.

Katzela, I. and M. Naghshineh. 1996. "Channel assignment schemes: a comprehensive survey," *IEEE Personal Communications*, 3, 3 (June), 10–31.

Koodli, R. and M. Punskari. 2001. "Supporting packet data QoS in next-generation cellular networks," *IEEE Communications Magazine*, 39, 2 (February), 180–188.

Kurose, J. F. and K. W. Ross. 2002. *Computer Networking*, Boston, MA, Addison-Wesley.

Kwok, M.-S. and H.-S. Wang. 1995. "Adjacent-cell interference analysis of reverse link in CDMA cellular radio system," IEEE PIMRC95, Toronto, Canada, September, 446–450.

Lee, W. C. Y. 1993. *Mobile Communications Design Fundamentals*, New York, Wiley-Interscience.

Lee, Y.-Z. *et al.* 2002. "An efficient and fair polling scheme for Bluetooth," *Proc. IEEE Milcom*, 1062–1068.

Levine, D. *et al.* 1997. "A resource estimation and admission control algorithm for wireless multimedia networks using the shadow cluster concept," *IEEE/ACM Transactions on Networking*, 5, 1 (February), 1–12.

Lin, Y.-B. 1997. "Reducing location update costs in a PCS network," *IEEE/ACM Transactions on Networking*, 5, 1 (February), 25–33.

Linnartz, J. P. 1993. *Narrowband Land-Mobile Radio Networks*, Boston, MA, Artech House.

Liu, Z. *et al.* 1996. "A demand-assignment access control for multi-code DS-CDMA wireless packet (ATM) networks," *Proc. IEEE Infocom'96*, San Francisco, CA, March, 713–721.

Lozano, A. and D. C. Cox. 1999. "Integrated dynamic channel assignment and power control in TDMA wireless communication systems," *IEEE Journal on Selected Areas in Communications*, 17, 11 (November), 2031–2040.

Lucky, R. W. 1965. "Automatic equalization for digital communication," *Bell System Technical Journal*, 44 (April), 558–570.

Mark, J. W. and W. Zhuang. 2003. *Wireless Communications and Networking*, Upper Saddle River, NJ, Pearson Education Inc.

Matsumoto, T. *et al.* 2003. "Overview and recent challenges of MIMO systems," *IEEE Vehicular Technology News*, 50, 2 (May), 4–9.

Meyerhoff, H. J. 1974. "Method for computing the optimum power balance in multibeam satellites," *Comsat Technology Review*, 4, 1 (Spring).

Michelson, A. M. and A. H. Levesque. 1985. *Error-Control Techniques for Digital Communication*, New York, John Wiley & Sons.

Miller, B. A. and C. Bisdikian. 2002. *Bluetooth Revealed*, 2nd edn, Upper Saddle River, NJ, Prentice-Hall PTR.

Milstein, L. 2000. "Wideband code division multiple access," *IEEE Journal on Selected Areas in Communications*, 18, 8 (August), 1344–1354.

Mitra, D. 1994. "An asynchronous distributed algorithm for power control in cellular systems," in *Wireless and Mobile Communications*, ed. J. M. Holtzman, Boston, MA, Kluwer/Academic.

Mitrou, N. M. *et al.* 1990. "A reservation multiple access protocol for microcellular mobile communication systems," *IEEE Transactions on Vehicular Technology*, 39, 4 (November), 340–351.

Moberg, J. *et al.* 2000. "Throughput of the WCDMA random access channel," *Proc. IST Mobile*

Rental, C. H. *et al.* 2002. "Comparative forward link traffic channel performance evaluation of HDR and 1EXTREME systems," *Proc. 55th IEEE Vehicular Technology Conference*, VTC 2002, Spring, 160–164.

Rose, C. and R. Yates. 1995. "Minimizing the average cost of paging under delay constraints," *Wireless Networks*, 1, 2 (July), 211–219.

Sarikaya, B. 2000. "Packet mode in wireless networks: overview of transition to third generation," *IEEE Communications Magazine*, 38, 9 (September), 164–172.

Schroeder, M. R. and B. S. Atal. 1985. "Code-excited linear prediction (CELP): High quality speech at very low bit rates," *Proc. IEEE Int. Conf. Acoust. Speech, Signal Process*, March, 937–950.

Schwartz, M. 1987. *Telecommunication Networks: Protocols, Modeling, and Analysis*, Reading, MA, Addison-Wesley.

Schwartz, M. 1990. *Information Transmission, Modulation, and Noise*, 4th edn, New York, McGraw-Hill.

Schwartz, M. 1996. *Broadband Integrated Networks*, Englewood Cliffs, NJ, Prentice-Hall.

Schwartz, M. and L. Shaw. 1975. *Signal Processing: Discrete Spectral Analysis, Detection, and Estimation*, New York, McGraw-Hill.

Schwartz, M., W. R. Bennett, and S. Stein. 1966. *Communication Systems and Techniques*, New York, McGraw-Hill; reprinted, IEEE Press, 1996.

Sköld, J. *et al.* 1995. "Performance and characteristics of GSM-based PCS," *Proc. 45th IEEE Vehicular Technology Conference*, 743–748.

Solana, A. H. *et al.* 2002. "Performance analysis of packet scheduling strategies for multimedia traffic in WCDMA," *Proc. 55th IEEE Vehicular Technology Conference*, VTC 2002, Spring, 155–159.

Steele, R. ed. 1992. *Mobile Radio Communications*, London, Pentech Press; New York, IEEE Press.

Stüber, G. L. 1996. *Principles of Mobile Communications*, Boston, MA, Kluwer Academic.

Stüber, G. L. 2001. *Principles of Mobile Communications*, 2nd edn, Boston, MA, Kluwer Academic.

Sutivong, A. and J. Peha. 1997. "Call admission control algorithms: proposal and comparison," *Proc. IEEE Globecom*.

TIA. 1992. EIA/TIA Interim standard, cellular system dual-mode mobile station–base station compatibility standard, IS-54-B, Telecommunications Industry Association.

Tiedmann, E. G., Jr. 2001. "cdma20001X: new capabilities for CDMA networks," *IEEE Vehicular Technology Society News*, 48, 4 (November), 4–12.

Varsou, A. C. and H. V. Poor. 2002. "Scheduling algorithms for downlink rate allocation in heterogeneous CDMA networks," *Journal of Communications and Networks*, 4, 3 (September), 199–208.

Verdu, S. 2002a. "Spectral efficiency in the wideband regime," *IEEE Transactions on Information Theory*, 48, 6 (June), 1319–1343.

Verdu, S. 2002b. "Recent results on the capacity of wideband channels in the low-power regime," *IEEE Wireless Communications*, 9, 4 (August), 40–45.

Viterbi, A. J. 1967. "Error bounds for convolutional codes and an asymptotically optimum decoding algorithm," *IEEE Transactions on Information Theory*, IT-13 (April), 260–269.

Viterbi, A. J. *et al.* 1994. "Soft handoff extends CDMA cell coverage and increases reverse link coverage," *IEEE Journal on Selected Areas in Communications*, 12, 8 (October), 1281–1288.

Viterbi, A. J. 1995. *CDMA, Principles of Spread Spectrum Communication*, Reading, MA, Addison-Wesley.

Vucetic, B. and J. Yuan. 2000. *Turbo Codes: Principles and Applications*, Boston, MA, Kluwer Academic.

Wang, J. Z. 1993. "A fully distributed location registration strategy for universal personal communications systems," *IEEE Journal on Selected Areas in Communications*, 11, 6 (August), 850–860.

WCDM. 2000. *WCDMA for UMTS*, ed. Harri Holma and Antti Toskala, John Wiley & Sons.

Willenegger, S. 2000. "cdma2000 physical layer: an overview," *Journal of Communications and Networks*, 2, 1 (March), 5–17.

Wozencroft, J. M. and I. M. Jacobs. 1965. *Principles of Communication Engineering*, New York, Wiley.

Xu, L. *et al.* 2002. "Dynamic bandwidth allocation with fair scheduling for WCDMA systems," *IEEE Wireless Communications*, 9, 2 (April), 26–32.

Yacoub, M. D. 1993. *Foundations of Mobile Radio Engineering*, Boca Raton, FL, CRC Press.

Yeung, K. L. and T.-S. P. Yum. 1995. "Cell group decoupling analysis of a dynamic allocation strategy in linear microcell radio systems," *IEEE Transactions on Communications*, 43, 2/3/4 (February–April), 1289–1292.

Yu, O. and V. Leung. 1997. "Adaptive resource allocation for prioritized call admission in ATM-based wireless PCN," *IEEE Journal on Selected Areas in Communications*, 15, 9 (September), 1208–1225.

Zander, J. 1992. "Distributed cochannel interference control in cellular radio systems," *IEEE Transactions on Vehicular Technology*, 41, 3 (August), 304–311.

Zhang, M. and T.-S. P. Yum. 1989. "Comparisons of channel assignment strategies in cellular telephone systems," *IEEE Transactions on Vehicular Technology*, 38, 4 (November), 211–218.

Zhang, H. 1995. "Service disciplines for guaranteed performance service in packet-switching networks," *Proc. IEEE*, 53 (October), 1374–1396.

Zorzi, M. and R. R. Rao. 1994. "Capture and retransmission control in mobile radio," *IEEE Journal on Selected Areas in Communications*, 12, 8 (October), 1289–1298.

Zürbes, S. *et al.* 2000. "Radio network performance of Bluetooth," IEEE International Conference on Communications, ICC, 1563–1567.

索　　引